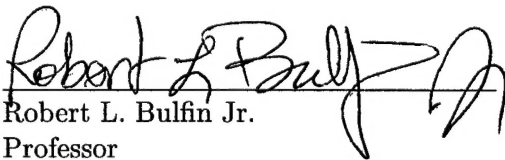


ADAPTIVE INVENTORY CONTROL FOR NON-STATIONARY DEMAND

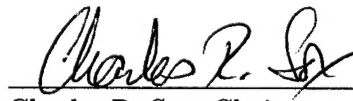
WITH PARTIAL INFORMATION

James Thomas Treharne

Certificate of Approval:



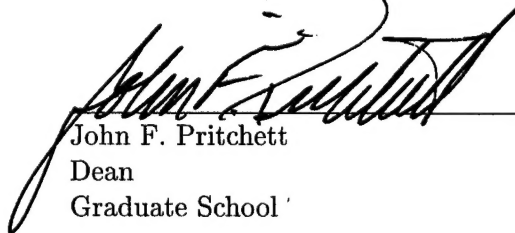
Robert L. Bulfin Jr.
Professor
Industrial and Systems Engineering



Charles R. Sox, Chair
Associate Professor
Industrial and Systems Engineering



Harriet Black Nembhard
Assistant Professor
Industrial and Systems Engineering



John F. Pritchett
Dean
Graduate School

ADAPTIVE INVENTORY CONTROL FOR NON-STATIONARY DEMAND
WITH PARTIAL INFORMATION

James Thomas Treharne

A Dissertation
Submitted to
the Graduate Faculty of
Auburn University
in Partial Fulfillment of the
Requirements for the
Degree of
Doctor of Philosophy

19990520 003

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Auburn, Alabama

June 11, 1999

DTIC QUALITY INSPECTED 4

VITA

James Thomas Treharne, son of Richard T. Treharne and Elizabeth J. Treharne, was born July 28, 1957 in Lakewood, Ohio. He attended Adlai E. Stevenson High School in Livonia, Michigan. He then graduated from the United States Military Academy in 1979 and was commissioned as a Regular Army 2nd Lieutenant in the United States Army Corps of Engineers. He was an Assistant Professor of Military Science at Auburn University from 1986-1990. In 1991 he earned a Master of Science in Industrial Engineering at Auburn University and was inducted into Phi Kappa Phi. His thesis was entitled, "RSSP- A Fortran Simulation Package For Use In Teaching Response Surface Methodology." He then served as a military operations research analyst with the Army Concepts Analysis Agency between 1991 and 1993. He is a 1994 graduate of the United States Army Command and General Staff College. LTC Treharne is married to the former Patricia A. Broaddus of Winchester, Kentucky. They have six children, Rebecca Joy, James David, Elizabeth Mae, Rachel Lynn, Julia Ann, and Lydia Grace.

DISSERTATION ABSTRACT

ADAPTIVE INVENTORY CONTROL FOR NON-STATIONARY DEMAND

WITH PARTIAL INFORMATION

James Thomas Treharne

Doctor of Philosophy, June 11, 1999

(M.S., Auburn University, 1991)

(B.S., United States Military Academy, 1979)

145 Typed Pages

Directed by Charles R. Sox

This dissertation presents optimal and suboptimal procedures to solve inventory control problems that have non-stationary demand and partial information. In each period, the underlying demand distribution may change according to a known Markov process. The problem is characterized as partial information because some parameter of the demand probability distribution is not known with certainty; however, there is a known prior distribution for the unknown parameter. In one case, there is a probability density function for the demand that has at least one unknown parameter, but this parameter has a known probability distribution. In another case, there is a set of candidate demand probability distributions. The parameter which indicates which demand is in effect at any given time is unknown, but has a known probability mass function. The control strategies are adaptive because the controllers learn information

about these unknown parameters over time and adapt accordingly. Because of the complexity of these problems, managers often estimate the unknown parameters and make decisions assuming the estimate is correct. The computational results presented in this dissertation demonstrate that there exist efficient and effective optimal and suboptimal procedures to solve these problems that potentially provide large cost savings compared with this current practice. The control strategies include open-loop feedback and limited look-ahead control for a finite horizon problem, which are compared to optimal and certainty equivalence control policies. A grid approximation and upper and lower bounds for an infinite horizon problem are also developed.

ACKNOWLEDGMENTS

I am grateful to Dr. Ed Unger and the ISE faculty for the opportunity to pursue this degree. I am especially thankful for my committee members, Dr. Charles Sox, Dr. Bob Bulfin and Dr. Harriet Nembhard. Dr. Sox has been a superb major professor in every respect. He has served as co-author of two papers which comprise Chapters 2 [39] and 3 [40] of this dissertation and will be co-author of a third paper coming from Chapter 4. This research has been supported in part by the National Science Foundation under grant number DMI 9813127.

I owe a debt of gratitude to my parents, Dick and Joan Treharne. I dedicate this degree to them. They made immense sacrifices to see that I had every opportunity to achieve personal and professional success. They taught me many values which serve me well today. Countless hours at ball games, wrestling meets, a long drive to buy my calculator (it did square-roots), pre-dawn drives along my paper route in the rain and snow, are just a few memories of my parent's love and devotion to me and my family.

I have been greatly blessed with a wonderful wife and six children. Patsy has always held our family together as I pursued professional and academic challenges. There is no advanced degree which recognizes all that she does. No man can have a more loving wife. I remain YGA.

Finally, I thank my Lord and Savior Jesus Christ for granting me more blessings and experiences than I ever deserved. To Him, I owe all.

Style manual or journal used conforms to that of the Institute of Industrial Engineers.

Computer software used is L^AT_EX 2_ε typesetting language with *auphd* style file developed at Auburn University.

TABLE OF CONTENTS

LIST OF FIGURES	x
LIST OF TABLES	xi
1 INTRODUCTION	1
2 REVIEW OF STATIONARY DEMAND AND PARTIAL INFORMATION	7
2.1 Introduction	7
2.2 Basic Model and Insights	10
2.3 Optimal Control	17
2.3.1 Uncensored Observations	17
2.3.2 Censored Observations	27
2.4 Suboptimal Control	31
2.4.1 Continuous Parameter	31
2.4.2 Discrete Parameter	36
2.5 Conclusion	40
2.A Optimality of (s, S) policy	44
3 POLICIES FOR NON-STATIONARY DEMAND AND PARTIAL INFORMATION	48
3.1 Introduction	48
3.2 Literature Review	49
3.3 Optimal Control Model and Insights	55
3.4 Suboptimal Control	63
3.4.1 Certainty Equivalence Control	64
3.4.2 Open-Loop Feedback Control	66
3.4.3 Limited Look-Ahead Control	67
3.5 Experiment Design	69
3.6 Results and Analysis	74
3.7 Conclusion	81
3.A Results	82
4 BOUNDS AND APPROXIMATIONS FOR ADAPTIVE INVENTORY CONTROL	91
4.1 Introduction	91
4.2 Literature Review	93
4.3 Optimal Control Model and Insights	98
4.4 Suboptimal Control	102
4.4.1 Grid Approximation	102

4.4.2	Upper Bound	108
4.4.3	Lower Bound	113
4.4.4	Myopic and Limited Look-Ahead Policies (Infinite Horizon)	115
4.5	Example Case	118
4.6	Conclusion	125
5	CONCLUSIONS AND FUTURE RESEARCH	127
	BIBLIOGRAPHY	130

LIST OF FIGURES

1.1	Non-Stationary Demand with Partial Information	5
2.1	Decision Process	11
2.2	Mixture Model	34
2.3	Cost Function with Fixed Order Cost	45
2.4	Possible Cost Function with Fixed Order Cost	46
3.1	Control Process	58
3.2	Discrete Demand Distributions	70
3.3	Conditional State Probability after Demand Observation	72
4.1	Control Process	100
4.2	Comparison of Stationary and Non-Stationary Bounds	114
4.3	Approximate Costs at π_1 for Different Spacings	119
4.4	Approximate Costs and Order Quantities at π_1	119
4.5	Cost-to-go function at various inventory levels	120
4.6	Initial Upper Bound	121
4.7	Upper Bound Improvement	122
4.8	Upper Bound Convergence	123

LIST OF TABLES

2.1	Notation	12
2.2	Distributions with State Space Reductions	24
3.1	Notation	56
3.2	Demand Distributions	69
3.3	Experiment Parameters	71
3.4	User Runtime	74
3.5	Percent of Cases Achieving Best Suboptimal Cost	75
3.6	Average Percent Above Optimal Cost	76
3.7	Summary of Results by Lead Time	76
3.8	Summary of Results by Critical Ratio	78
3.9	Summary of Results by Transition Matrix	78
3.10	Summary of Results by the Prior	80
3.11	Results: Lead time=0; Trans Matrices 1, 2 & 3	82
3.12	Results: Lead time=0; Trans Matrices 4 & 5	83
3.13	Results: Lead time=0; Trans Matrices 6 & 7	84
3.14	Results: Lead time=1; Trans Matrices 1, 2 & 3	85
3.15	Results: Lead time=1; Trans Matrices 4 & 5	86
3.16	Results: Lead time=1; Trans Matrices 6 & 7	87
3.17	Results: Lead time=2; Trans Matrices 1, 2 & 3	88

3.18 Results: Lead time=2; Trans Matrices 4 & 5	89
3.19 Results: Lead time=2; Trans Matrices 6 & 7	90
4.1 Notation	99
4.2 Demand Distributions	118
4.3 Experiment Parameters	118
4.4 Initial Subgradients and Actions	124
4.5 Lower Bound Run Times and Percent Deviation from UB	124
4.6 Upper Bound Run Times and Percent Deviation from LB	125

CHAPTER 1

INTRODUCTION

The purpose of this dissertation is to investigate a fundamental inventory control problem in which the demand distribution is non-stationary and the information about the distribution is not complete. We title this dissertation “Adaptive Inventory Control for Non-Stationary Demand with Partial Information” because we are interested in examining such problems in which we receive state-dependent information about the underlying, but unknown demand distribution. The system is adaptive because each period information is learned about the value of the unknown parameter. The controller uses this information as it accumulates to adjust future decisions accordingly. Therefore, stocking decisions are made under the assumption that new information dependent on the current state will be observed in each time period, and that this new information will be incorporated into future decisions. Such an adaptive control policy is important because the demand distribution may change each period. The current business environment is marked by two important characteristics. The first is that demand in today’s markets is often highly volatile. This volatility is due to a number of factors, including the proliferation of products, the global marketplace, shortened product life-cycles, and the rapid growth of technology. Manufacturers may no longer safely assume that product demand will remain stationary for long periods of time. The second characteristic is that we are now inundated with vast amounts of information. Despite this

wealth of information, managers still face a great deal of uncertainty about demand. At the same time, we now have the computer capability to solve problems with greater degrees of complexity. Thus, we must be able to use all available information to make the best inventory decision, knowing that in the next period circumstances will be different and new information will be available. Managers often disregard uncertainty by using point forecasts as if they are accurate. Because of the advances of computational speed and the results reported in this dissertation, such simplifying assumptions are no longer necessary for efficient and effective inventory control decisions. Our goal is to show that there are effective and efficient procedures to solve problems when the demand distribution is non-stationary and only partial information is available. Generally, inventory problems can be divided into four classes. The first class is stationary demand with complete information. The second class is non-stationary demand with complete information. The third class is stationary demand with partial information. The final class is non-stationary demand with partial information. Problems with complete information about the current demand state have been studied fairly comprehensively. The introduction of partial information significantly increases the difficulty of the problem. By partial information we mean that the demand distribution possesses an unknown parameter that may be either discrete or continuous. Some information in the form of a prior distribution is known about the unknown parameter. Very little work has been done when the demand distribution is non-stationary and also has partial information about the current distribution.

To achieve our stated goal we present in Chapter 2 a comprehensive review of adaptive inventory control for stationary demand. We provide a basic model that can represent many realistic problems in this problem class. Currently, this class of problems is not well addressed in the literature. We show two approaches for this type of problem. In the first approach, the demand distribution is assumed to belong to a known family but have at least one unknown parameter, which itself has a known probability distribution which is called the prior distribution. The second approach is to assume that there is a discrete set of possible demand distributions and a prior belief for which distribution is in effect. In Chapter 2 we present both a model as well as many of its characteristics. We also present some results for optimal strategies and their properties. In many cases, suboptimal solutions will be necessary, and we discuss several such suboptimal control strategies. We discuss such key concepts as sufficient statistics, conjugate families, censored versus uncensored observations, and state space reduction.

In Chapter 3 we extend our model to the case of non-stationary demand with partial information. Very little work has been done in this important area. Managers, especially in these cases, will often estimate or forecast the unknown parameters and assume they are completely known. By doing so, they are ignoring a significant amount of the uncertainty in the system, and are potentially making costly decisions. We show in this chapter that there are many other efficient suboptimal control policies that provide excellent results by incorporating more of the inherent uncertainty into the decision process. Fully considering uncertainty, even over a limited horizon, may lead to solutions that are

significantly better than certainty equivalence control methods. In both Chapters 3 and 4 we model the process as a composite-state, partially observed Markov decision process.

In Chapter 4 we continue to study the non-stationary problem with partial information, and we present valid techniques to compute bounds and approximations for this problem. In this chapter, we assume that the problem is one with a discounted, infinite horizon because many practical problems can be accurately modeled as infinite horizon problems. Again, we model the demand process as a composite-state, partially observed Markov decision process, and we extend known techniques for the stationary problem. Most problems do not have a tractable optimal solution because the probability state space is uncountably infinite. However, we can accurately approximate the optimal cost function for some problems using a grid approximation. The approximation is asymptotically exact as the grid spacing approaches zero. However, a small spacing can be computationally infeasible. Because we show that the cost function is piece-wise linear and concave in the probability state space, we can compute subgradients that form an upper bound. We can easily calculate these subgradients at each extreme point, and we can then examine the problem at other values in the probability state space and iteratively add subgradients that improve the upper bound. It is also possible to calculate a lower bound to this function. We show both how to calculate this lower bound and how to improve the bound in a similar way. Finally, we demonstrate some of these techniques with an example.

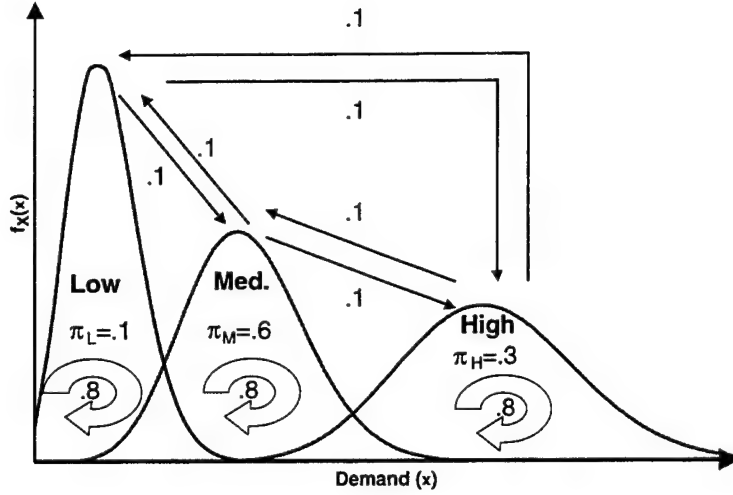


Figure 1.1: Non-Stationary Demand with Partial Information

In Figure 1.1 we present an example of the situation that we address. In this case the demand distribution belongs to one of three fully defined distributions. The demand distribution currently in effect is unknown, although π_L , π_M , and π_H represent the beliefs that the true distribution is either the low, medium, or high distribution. Because the true distribution is unknown, we have partial information available. The figure also shows that in the next period the demand distribution may shift according to a Markov process. In this case, the probability that the demand does not change in the next period is 0.8 (broad arrows), while the probability that it moves to one of the other two distributions is 0.1 each (thin arrows). The objective of this problem is to minimize the total costs over the given horizon, given the non-stationary demand process and partial information about its state.

In summary, the following chapters demonstrate viable solution procedures to solve problems with partial information and either stationary or non-stationary demand. These techniques can potentially provide substantial cost savings over procedures in common practice today.

CHAPTER 2

REVIEW OF STATIONARY DEMAND AND PARTIAL INFORMATION

2.1 Introduction

This chapter provides inventory managers valuable insights and theory for determining control policies when faced with partial information about the demand process. More specifically, the chapter reviews key literature for the problem of controlling the inventory of a product with partially observed, stationary, random demand. Practitioners usually presume it is necessary to model their problem as a “complete information” problem in order to solve a realistic problem. However, this chapter demonstrates that it may be both practical and worthwhile to solve the “partial information” problem. The demand process is not completely observed but is only partially observed through the random demand observations. Because the control decisions are made with only partial information about the demand process, the level of uncertainty and the cost of suboptimal decisions is much higher than for most inventory problems considered in the research literature. This problem is an accurate representation of the inventory control problems faced by many organizations. However, it has not been widely addressed in the inventory literature or by existing decision support systems; therefore, inventory managers are forced to make potentially costly simplifying assumptions when addressing this challenging problem. This chapter discusses the existing research for this problem and points out its critical components and areas for future research. Although we do

not provide an exhaustive summary of current research, we do present some of the most critical existing research for the problem.

Many inventory control problems are significantly affected by a lack of demand information. We focus our attention in this chapter to that part of the life cycle in which demand is assumed stationary. As product life cycles continue to shrink, a larger fraction of that life cycle lies in the initial and terminal periods where there is a great deal of uncertainty in the demand process. Some of this uncertainty applies also to the stationary stage of the life cycle. Consequently, the area of our concern in which demand is considered stationary will be marked by less precise information about its unknown parameters. Also, the proliferation of products continues to dilute the historically stable demand for individual products, leading to more volatile and uncertain demand. This proliferation also increases the need for automated or semi-automated inventory control systems. The competition in some markets for commodity-like products also creates more volatile, uncertain demand than is assumed in most conventional models. This research also has a potential for broader impact on other problems that involve partially observed, random processes. Such problems arise in a variety of problem domains such as supply chain planning and control (Lee *et al.* [22]), maintenance planning, yield planning and management, military materiel management, and process control.

There are two main approaches for modeling partial information inventory control problems. In the first approach, the demand distribution is assumed to belong to a known distribution family, but at least one parameter is unknown. A known *a priori*

distribution is assumed for the unknown parameter, and subsequent demand observations are used to compute a *posterior* distribution for this parameter. In the second approach, the demand distribution is assumed to belong to a known and discrete set of candidate distributions. A known, discrete *a priori* distribution is assumed for this set which is also used along with subsequent observations to compute a *posterior* distribution. Both types of models can be solved optimally using stochastic dynamic programming for only small problems.

Because optimal control requires restrictive assumptions or is computationally intractable, most practical solutions rely on suboptimal control. A very common suboptimal control strategy is a type of certainty equivalence controller. In this type of policy, the parameter estimates are generated from historical data and then are assumed to be *known with certainty* in the control phase. For example, they might assume a Poisson distribution with a known parameter based on last year's demand. The parameter estimates may be revised periodically, but this approach takes no further consideration of the inherent uncertainty in those estimates. A very important area of investigation is to determine when this class of strategies is effective and what other types of strategies can be efficiently employed when they are ineffective.

The chapter is organized as follows. In Section 2.2 we present a basic model for this problem and discuss some of its characteristics. The next section, Section 2.3, presents some results for optimal control strategies and their properties. Section 2.4 discusses

several suboptimal control strategies, and the last section provides some concluding remarks and areas for valuable future research.

2.2 Basic Model and Insights

We first define the basic inventory model and discuss some of its important characteristics. Any extensions or exceptions to this model are indicated where they occur in the chapter. The problem is stationary, that is, the demand is independent and identically distributed across periods, and the various cost parameters remain constant. Thus, our focus is only on that part of the life cycle that is “stationary”. However, the decision maker does not have complete information about the parameters of the demand distribution. At least one parameter, τ , is unknown, and this unknown parameter has a known *a priori* distribution $f(\tau)$. Although we use notation that presumes f is continuous, the model is easily extended to cases where f is discontinuous or discrete. The planning horizon may be finite or infinite. Stockouts are either backordered or become lost sales. If stockouts are lost, the lost sales may be observed or unobserved. If lost sales are observed, all demand observations are exact and the starting inventory for the next period is at least zero. The order lead time is a known constant, possibly zero. The procurement cost may include a known fixed cost. The procurement, shortage and holding costs may be linear, convex, or non-linear functions. The current stock level is always known with certainty, and the salvage values at the end of the horizon may be positive or negative. Stock cannot be freely disposed of although some authors allow

this primarily to simplify the solution. Finally, stock does not deteriorate, although this can be easily incorporated into the model when it is a fixed proportion of on-hand inventory. These assumptions make the solution to the general problem quite difficult to solve. However, in many special cases or under relaxed assumptions there are very practical solutions with useful results.

Figure (2.1) describes the decision process. The decision maker uses current information to choose an action, the order quantity. An order that was placed a lead time in the past is then received. Demand is then observed. In the case of unobserved lost sales, only actual sales are observed. The period costs are then assessed. The decision maker records demand or sales data and updates the available information and repeats the process.

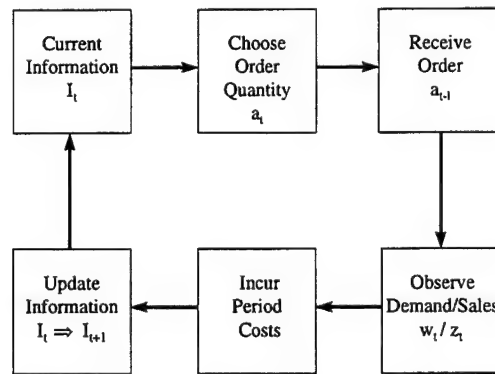


Figure 2.1: Decision Process

Table 2.1 shows the notation that is used throughout the chapter. Additional terms are defined as needed.

<i>Var</i>	<i>Description</i>	<i>Var</i>	<i>Description</i>
T	length of planning horizon	α	discount rate
$p(\cdot)$	stockout penalty cost	l	positive lead time
$h(\cdot)$	holding cost	S_t	sufficient statistic in period t
$c(\cdot)$	procurement cost	I_t	vector of information in period t
K	positive fixed order cost	τ	unknown parameter
x_t	inventory level in period t	$f_t(\tau I_t)$	prior distribution of τ in t
u_t	inventory position	$\phi(w S_t)$	conditional demand in t
a_t	order quantity in period t	$\bar{a}_{t,l}$	vector of orders in period t
w_t	demand in period t $\phi(w; \tau)$	$\phi(w; \tau)$	demand density
z_t	sales in period t		

Table 2.1: Notation

The information available at the beginning of period t , I_t , can often be summarized by a sufficient statistic, S_t . Such a statistic contains the necessary information to compute an optimal policy for period t . The exact form of this statistic depends on the specific problem instance through the prior and demand distributions. A statistic is *sufficient* if given a prior distribution on the parameter τ , $f(\tau)$, the posterior distribution on the parameter τ , $f_t(\tau|S_t)$, is a function of the sufficient statistic only, and not the individual sample observations. Therefore, to update the prior distribution of a parameter to a posterior distribution on the same parameter, only the value of the sufficient statistic needs to be known. DeGroot [8] discusses sufficient statistics in detail and provides numerous examples. The initial prior distribution is $f(\tau) = f_0(\tau|I_0)$. The current prior, $f_t(\tau|I_t)$, is always a sufficient statistic that is easily revised using Bayes' formula

$$f_{t+1}(\tau|I_{t+1}) = \frac{\phi(w_t; \tau) f_t(\tau|I_t)}{\int_0^\infty \phi(w_t; \xi) f_t(\xi|I_t) d\xi}. \quad (2.1)$$

The conditional demand distribution depends on the current prior through the equation

$$\phi_t(w|S_t) = \int_0^\infty \phi(w; \tau) f_t(\tau|I_t) d\tau. \quad (2.2)$$

We now present our general model formulation. We first present a basic model and show how we can modify it to account for positive lead time, fixed order costs, and unobserved lost sales. When lead time is zero with backordering and no fixed costs, the control action in period t raises the inventory level to a value y with an expected holding and backorder cost

$$L_t(y|S_t) = \int_0^y h(y-w)\phi_t(w|S_t)dw + \int_y^\infty p(w-y)\phi_t(w|S_t)dw. \quad (2.3)$$

The state transition equation for the inventory level is $x_{t+1} = x_t + a_t - w_t$, and the cost-to-go function is then

$$J_t^T(x_t, S_t) = \inf_{y \geq x_t} \{c(y - x_t) + L_t(y|S_t) + \alpha \int_0^\infty J_{t+1}^T(y - w, S_{t+1})\phi_t(w|S_t)dw\}. \quad (2.4)$$

This function calculates the total discounted costs from time t to the end of the horizon, T . When we account for positive lead time, we augment the state space to include the outstanding orders at time t . Define the vector $\bar{a}_{t,l} = (a_{t-l}, a_{t-l+1}, \dots, a_{t-1})$, then the state vector at time t is $(x_t, S_t, \bar{a}_{t,l})$. The inventory transition equation is $x_{t+1} = x_t + a_{t-l} - w_t$, and the order transition equation is $\bar{a}_{t+1,l} = \bar{a}_{t,l} - \{a_{t-l}\} + \{a_t\}$. For an

arbitrary inventory level v , define the single period expected inventory cost function as

$$L_t(v|S_t, l) = \int_0^v h(v - w_{t,l})\phi_t(w_{t,l}|S_t, l)dw_{t,l} + \int_v^\infty p(w_{t,l} - v)\phi_t(w_{t,l}|S_t, l)dw_{t,l}, \quad (2.5)$$

where $w_{t,l} = \sum_{k=0}^l w_{t+k}$. This cost function accounts for the positive lead time in computing the expected inventory costs for the inventory level at time $t + l$. The cost-to-go function with positive lead times becomes

$$J_t^T(x_t, S_t, \bar{a}_{t,l}) = \inf_{a_t \geq 0} \{c(a_t) + L_t(x_t + \sum_{k=0}^l a_{t-k}|S_t, l) + \alpha \int_0^\infty J_{t+1}^T(x_t + a_{t-l} - w_t, S_{t+1}, \bar{a}_{t+1,l})\phi_t(w|S_t)dw\}. \quad (2.6)$$

This problem can be simplified by formulating it in terms of the amount of inventory on-hand and on-order at the beginning of t , $u_t = x_t + \sum_{k=1}^l a_{t-k}$, which is referred to as the *inventory position*. The inventory position follows the transition equation $u_{t+1} = u_t + a_t - w_t$, and the DP recursion becomes

$$J_t^T(u_t, S_t) = \inf_{a_t \geq 0} \{c(a_t) + L_t(u_t + a_t|S_t, l) + \alpha \int_0^\infty J_{t+1}^T(u_t + a_t - w_t, S_{t+1})\phi_t(w|S_t)dw\}. \quad (2.7)$$

Including lead time will significantly increase the effect of uncertainty in the problem due to the uncertainty of demand over the lead time. Adding a fixed cost to the problem only requires a slight modification to the model at this point. The cost-to-go functions

becomes

$$J_t^T(u_t, S_t) = \inf_{a_t \geq 0} \{K \cdot 1\{a_t \geq 1\} + c(a_t) + L_t(u_t + a_t | S_t, l) + \alpha \int_0^\infty J_{t+1}^T(u_t + a_t - w_t, S_{t+1}) \phi_t(w | S_t) dw\}, \quad (2.8)$$

where $1\{\cdot\}$ is an indicator function. The final addition to our model is that of unobserved lost sales. The primary distinction from this additional assumption is that the measurement process is less precise. When stockouts are backordered, the demand, $\{w_t\}$, is the measurement process of the underlying process because the demands are directly observed. We assume that when stockouts are lost that only actual sales are observed so that $\{z_t\}$ where $z_t = \min\{x_t + a_{t-l}, w_t\}$ is the measurement process. In this case, z_t is the actual sales in period t as opposed to the demand w_t . The unobserved lost sales case results in censored observations and increases the value of on-hand stock as a means of more accurately observing demand. In lost sales cases, the penalty cost must also account for lost profit per unit of lost sales. The information vector and sufficient statistic may be more difficult to calculate due to the unobserved lost sales. This leads to the final form for the cost-to-go function in our model,

$$J_t^T(x_t, S_t, \bar{a}_{t,l}) = \inf_{a_t \geq 0} \{K \cdot 1\{a_t > 1\} + c(a_t) + L_t(x_t + a_{t-l} | S_t) + \alpha \int_0^\infty J_{t+1}^T(x_t + a_{t-l} - z_t, S_{t+1}, \bar{a}_{t+1,l}) \phi_t(z | S_t) dz\}. \quad (2.9)$$

Also, the cost-to-go function must include the vector of orders, $\bar{a}_{t,l}$, because the possibility of lost sales precludes the simplification from using the inventory position as a state variable. This greatly complicates computation for the lost sales case with positive lead time.

There are some very important modeling assumptions that arise in this basic formulation. The most critical assumptions are the choice of an *a priori* distribution, f , for the unknown parameter τ and the choice of a distribution for demand, $\phi(w; \tau)$. These choices are particularly important when τ is continuous because they affect the tractability of the problem. If the *a priori* and demand densities are chosen from a pair of conjugate families, then the Bayesian calculation of the *posterior* distribution of τ is greatly simplified; otherwise, this calculation requires numerical integration. Furthermore, the cost-to-go function in Equation (2.4) has two state variables, x_t and S_t . An even more selective choice of densities can lead to a reduction of the state space to a single dimension which further improves the tractability of the model. The drawback to these assumptions is that they may lead to a model that does not accurately approximate the distribution for the actual problem. In these cases, τ may be modeled as a discrete parameter with finite range, the calculation of the *posterior* distribution is simplified and does not require the use of conjugate families. However, there are no known results that reduce the size of the state space for the discrete problem. This situation leads to an opportunity for an interesting line of research which is to investigate the effectiveness of developing a near-optimal control policy for a discretized approximation of the true prior f .

There are also some interesting and useful properties which are known for this problem. The policy structure is similar to that of the fully observed problem except that it is dependent on the current prior, $f_t(\tau)$. In the case of no fixed ordering cost with linear purchase, holding, and penalty costs, the policy is a state-dependent base stock control policy, $S(f_t(\tau))$, see Scarf [33]. In the case of a positive fixed order cost the policy is state-dependent with two critical numbers, $(s(f_t), S(f_t))$, see Lovejoy [28]. These policy structures reduce the computational requirements and are an intuitive extension of the policies for the fully observed problems. Another important property is that in the case of back ordering the optimal policy converges to the maximum likelihood policy, *i.e.*, the MLE of τ , $\hat{\tau}$, is assumed to be the known value of τ , see Scarf [33]. This means that once enough information is gathered about the value of τ , then the state-dependent policy can be replaced by the certainty equivalent policy using $\hat{\tau}$ without a significant loss of optimality. Thus, an interesting research question is raised, which is to determine when this policy transition should take place.

2.3 Optimal Control

2.3.1 Uncensored Observations

Many practitioners assume, although sometimes incorrectly, that their demand is observed exactly. When this is the case, the solution to the problem is somewhat simplified. The following section discusses uncensored observation cases.

Basic Properties

Several authors consider the case of full backordering. Similar results would hold in lost sales cases in which demand, although lost, is still fully observed. We first discuss some general properties of the problem from three papers by Scarf [33], Karlin [17], and Iglehart [15]. We then provide results from Scarf [34] and Azoury [4] that are not only more specific in nature, but also provide a reduction in the state space for the problem.

Herbert Scarf [33] was a pioneer in the use of Bayesian techniques for inventory control. Scarf assumes that holding, penalty, and procurement costs are all linear. There are no fixed order costs, no end of horizon salvage value, lead time is zero, demand is backordered, and future costs are discounted. He also assumed that the demand density was in the exponential class: $\{\phi(w; \tau) = \beta(\tau)e^{-w\tau}r(w)\}$ where $r(w) = 0$ for $w < 0$. The exponential class of distributions (including Gamma, Poisson, and negative binomial) is particularly convenient because it has a scalar sufficient statistic, $S_t = \sum_{j=1}^{t-1} w_j$.

Scarf's first key result is that optimal policies for this problem are defined by critical numbers, $\bar{x}_t(S_t)$, *i.e.*, base stock control. Scarf proves that $\bar{x}_t(S_t)$ is the unique solution to Equation (2.4), and that it is monotone increasing in S_t . Additionally, when demand is sufficiently small in t , then no order is placed in the following period, $t + 1$. Analogous results hold for the infinite horizon version.

Since critical numbers are difficult to obtain in general, a very important result is that the critical numbers, $\bar{x}_t(S_t)$, have an asymptotic relationship to the critical numbers in the case that the demand distribution is fully known with mean equal to S_t . Define

this critical number as $\bar{X}(S_t)$, then $\bar{x}_t(S_t) \sim \bar{X}(S_t) + \frac{a(S_t)}{t}$. Scarf also provides the explicit equation for $a(S_t)$, which may have a positive or negative value. This asymptotic relationship implies that the Bayesian stock level converges to the maximum likelihood policy. So, for sufficiently large t , the Bayesian stock level is readily approximated.

Karlin [17] first analyzes a “non-stationary” problem in which the distribution can vary each period. Using those results, he then considers our case of stationary demand in which the demand is not fully known. Karlin examines two cost structures, one with linear purchase costs and another with convex purchase costs. In each case, holding and penalty costs are convex. Karlin asserts that similar results hold when there are positive lead times. He also allows both backordering and observable lost sales. He extends the results of Scarf (exponential family) to include the range family and provides some ordering properties. As in Scarf, there is an *a priori* distribution on the unknown demand parameter. Distributions in the range family are of the form $\phi(w; \tau) = \gamma(\tau)q(w)\psi(w; \tau)$ where $\psi(w; \tau) = 1$ for $0 \leq w \leq \tau$, and 0 otherwise. The range family includes all continuous, positive random variables with a bounded density that is truncated at some positive number τ , e.g., uniform. The sufficient statistic for this family is $S_t = \max_{1 \leq i \leq t} w_i$. Karlin demonstrates two stochastic orderings for both the exponential and range families. Karlin shows that for a given time, t , a smaller sufficient statistic will imply that the conditional demand density is stochastically smaller than the density associated with the larger statistic: If $S_t < S'_t$, then $\phi_t(\cdot | S_t)$ is stochastically smaller than $\phi_t(\cdot | S'_t)$ (written

$\phi_t(\cdot|S_t) \subset \phi_t(\cdot|S'_t)$). For a fixed value of a sufficient statistic ($S = S_{t+1} = S_t$), the conditional demand density at $t+1$ is smaller than the density at time t : $\phi_{t+1}(\cdot|S) < \phi_t(\cdot|S)$. With linear purchase costs and other costs convex, the optimal policy is a single critical number, $\bar{x}_t(S_t)$. If the purchase costs are generalized to be strictly convex with linear shortage costs and convex holding costs, the policy is a function of two numbers. The policy is of the (s, S) type and is to order up to $y(x; \phi_1, \phi_2, \dots)$ if $x < \bar{x}(\phi_1, \phi_2, \dots)$; otherwise, do not order. This leads to the result that for a one-period model in which there is a sufficient statistic S_t or S'_t where $(S_t < S'_t)$, the order-up-to quantities for the two respective cost structures are such that $\bar{x}_1(S_t) \leq \bar{x}_1(S'_t)$ and $\bar{y}_1(x|S_t) \leq \bar{y}_1(x|S'_t)$ for all $x \geq 0$. Then, by induction, where again $(S_t < S'_t)$, the initial order-up-to quantities for the T period problem are such that $\bar{x}_T(S_t) \leq \bar{x}_T(S'_t)$ and $\bar{y}_T(x|S_t) \leq \bar{y}_T(x|S'_t)$ for all $x \geq 0$.

Iglehart [15] extends the work of both Scarf and Karlin. He extends Scarf by including the range family as well as the exponential family in his results. The shortage and penalty costs are convex while procurement costs remain linear. Iglehart assumes full backordering and zero lead time. However, similar results still hold when either assumption is relaxed. As in Scarf and Karlin, there is an unknown parameter for the demand which has a known *a priori* distribution. Iglehart also presents a more complete set of theoretical properties and proofs than Karlin. He first demonstrates that the optimal policy in Equation (2.4) is represented by a single critical number, $\bar{x}_T(S_t)$. The cost function, $J_t^T(x_t, S_t)$, is convex with respect to x . He then shows that for

a fixed S_t that $\bar{x}_T(S_t) \leq \bar{x}_{T+1}(S_t)$ and that $\bar{x}_T(S_t) \leq \hat{x}(S_t)$ where $\hat{x}(S_t)$ is smallest root of the equation $c(1 - \alpha) + L'(y|S_t) = 0$. Also, $\lim_{T \rightarrow \infty} J^T(x|S_t) = J(x|S_t)$. He presents more complete results than Karlin about how the critical numbers vary with changes in the information. He shows that for $S_t < S'_t$, then $\bar{x}_T(S_t) \leq \bar{x}_T(S'_t)$ and $\frac{\partial}{\partial x} J^T(x|S_t) \geq \frac{\partial}{\partial x} J^T(x|S'_t)$. That is, a larger sufficient statistic at time t will lead to an equal or larger base stock level for the T -period horizon problem. Also, the slope of the cost function will be larger at any point x for the function with the smaller sufficient statistic. He proves the following results for all fixed values of S and x . The critical number $\bar{x}_T(S, t + 1) \leq \bar{x}_T(S, t)$ while $\frac{\partial}{\partial x} J^T(x|S, t + 1) \geq \frac{\partial}{\partial x} J^T(x|S, t)$. Iglehart shows the asymptotic behaviors of the conditional demand when the sufficient statistics are fixed for large t for both exponential and range cases. For example, if the maximum likelihood estimator of τ is $\hat{\tau}_t$, and if $f(\hat{\tau}_t) > 0$ and $f(\tau)$ is continuous in the neighborhood of $\hat{\tau}_t$, then $\lim_{t \rightarrow \infty} \phi(w|S_t) = \phi(w; \hat{\tau}_t)$. Scarf showed that $\lim_{t \rightarrow \infty} J(x, S_t) = J(x, \hat{\tau}_t)$ and $\lim_{t \rightarrow \infty} \bar{x}(S_t) = \bar{x}(\hat{\tau}_t)$. He concludes by showing the following asymptotic property. If $f(\hat{\tau}_t) > 0$ and continuous, then $\bar{x}(S_t) \sim \bar{x}(\hat{\tau}_t) + a(\hat{\tau}_t)$ as $t \rightarrow \infty$ where $a(\hat{\tau}_t)$ is defined appropriately for both the exponential and range families.

State Space Reduction

We now focus attention on results that greatly simplify the calculation of optimal policies. In a subsequent paper [34], Scarf shows that in some instances the critical number can be determined as a function of one variable. He first assumes that procurement,

holding, and penalty costs are all linear. Lead time is zero, but he notes that his procedures can be adapted for use if there is positive lead time and backordering. He also assumes that the demand distribution is in the gamma family: $\phi(w; \tau) = \frac{\tau(\tau w)^{a-1} e^{-\tau w}}{\Gamma(a)}$, where a is a known constant and τ is an unknown parameter. The *a priori* distribution is also gamma with known parameters, λ and b , so that $f(\tau) = \frac{\lambda(\lambda \tau)^{b-1} e^{-\lambda \tau}}{\Gamma(b)}$. In this case the sufficient statistic is $S_t = \sum_{j=1}^{t-1} w_j$. He first demonstrates the relationship that reduces the problem to one of a single state variable. The conditional demand distribution can be reformulated as

$$\phi_t(w|S_t) = \frac{1}{(S_t + \lambda)} \Phi_t \left(\frac{w}{S_t + \lambda} \right), \quad (2.10)$$

$$\text{where } \Phi_t(w) = \frac{\Gamma(ta + b)}{\Gamma(a)\Gamma[(t-1)a + b]} \frac{w^{a-1}}{(1+w)^{ta+b}}. \quad (2.11)$$

He then demonstrates that the cost function can be computed as a scalar multiple of a related single state variable problem, *i.e.*, $J_t(x, S_t) = (S_t + \lambda) \tilde{J}_t(\frac{x}{S_t + \lambda})$,

$$\text{where } \tilde{J}_t(\tilde{x}) = \min_{y \geq \tilde{x}} \left\{ c(y - \tilde{x}) + L_t(y) + \alpha \int_0^\infty \tilde{J}_{t+1} \left(\frac{y-z}{1+z} \right) (1+z) \phi_t(z) dz \right\}. \quad (2.12)$$

He also shows that the critical numbers are proportional to $\lambda + S_t$. Because of the state space reduction in this special case, the solution to the inventory problem with incomplete information is no more difficult than a problem with complete information.

This reduction in the state space should be very attractive to practitioners facing these types of problems.

Azoury [4] extended Scarf's work by showing more general conditions for a reduction in the state space of the Bayesian model. Azoury assumed a finite horizon with backordering and that the ordering, holding, and shortage costs are linear and lead time is zero. Azoury considers both depletive and nondepletive (repairable) inventory models. She chooses a series of demand distributions that have a prior which comes from a conjugate family. This ensures that the posterior distribution on the unknown parameter is in the same family as the prior. See DeGroot [8] for a detailed description of conjugate families.

We will examine results of the depletive model. Once again the model is defined as in Equation (2.4). The two conditions that allow the state space reduction are:

1. There exists a set of functions $q_t(S_t)$ such that $\phi_t(w|S_t) = \left(\frac{1}{q_t(S_t)}\right) \psi_t\left(\frac{w}{q_t(S_t)}\right)$ where $\psi_t(\cdot)$ is a probability density function of scaled demand that depends only on t .
2. $q_{t+1}(S_{t+1}) = q_t(S_t)U_{t+1}\left(\frac{w}{q_t(S_t)}\right)$ for some continuous real valued function U_{t+1} such that

$$\int_0^\infty U_{t+1}(z)\psi_t(z)dz < \infty.$$

<i>Demand</i>	<i>Prior</i>	<i>Suff. Stat.</i>	$\bar{x}_t(S_t)$
Uniform $\phi(w; \tau) = \begin{cases} 1/\tau & \text{if } 0 \leq w \leq \tau \\ 0 & \text{otherwise.} \end{cases}$	Pareto $g(\tau) = aR^a/\tau^{a+1}$ $\tau > R; a, R \text{ known}$	$\max_{1 \leq i \leq t-1} w_i$	$\gamma_t \max(S_t, R)$
Weibull $\phi(w; \tau) = \tau k w^{k-1} e^{-\tau w^k}$ $w \geq 0, k \text{ known}$	Gamma $g(\tau) = \frac{b^a \tau^{a-1} e^{-b\tau}}{\Gamma(a)}$ $\tau \geq 0; a, b > 0$	$\sum_{i=1}^{t-1} w_i^k$	$\gamma_t (b + S_t)^{1/k}$
Periodic/Gamma $w_t = k_t Z_t$ k_t is periodic index $\phi(z; \tau) = \frac{\tau^\lambda z^{\lambda-1} e^{-\tau z}}{\Gamma(\lambda)}$ $z \geq 0; \lambda, \tau > 0$	Gamma $g(\tau) = \frac{b^a \tau^{a-1} e^{-b\tau}}{\Gamma(a)}$ $\tau \geq 0; a, b, > 0$	$\sum_{i=1}^{t-1} (w_i/k_i)$	$\gamma_t k_t (b + S_t)$

Table 2.2: Distributions with State Space Reductions

The equation of the single variable is

$$V_t(\tilde{x}) = \min_{y \geq \tilde{x}} \{c(y - \tilde{x}) + H_t(y) + \alpha \int_0^\infty V_{t+1} \left(\frac{y - z}{U_{t+1}(z)} \right) U_{t+1}(z) \psi_t(z) dz\}, \quad (2.13)$$

where $H_t(y) = h \int_0^y (y - z) \psi_t(z) dz + p \int_y^\infty (z - y) \psi_t(z) dz$ for $y \geq 0$. The function $V_t(\cdot)$ is convex, hence a base stock policy is optimal. Let γ_t achieve the minimum of $V_t(\tilde{x})$ in (2.13). Azoury then proves for the original problem that $\bar{x}_t(S_t) = \gamma_t q_t(S_t)$. Table 2.2 shows three sets of distributions that satisfy the conditions. Note that for $k_t = 1$, the third example is the result due to Scarf [34]. The last column shows the optimal order-up-to level for the finite horizon model and is a function of the sufficient statistic. The critical level is always proportional to γ_t , the value of y that minimizes Equation (2.13).

Azoury's results for the nondepletive models are analogous to those of the depletive models. Additionally, the results of the depletive model also hold for positive lead time. When Azoury or Scarf's assumptions hold, managers can often solve partial information problems with relative ease.

Scarf and Azoury showed how under certain conditions these problems can be reduced to single dimensional state space. Lovejoy [25] adds two more assumptions to those of Azoury and shows under these additional assumptions that policies can be reduced to a zero dimensional state space and solved by a simple critical fractile policy. For problems that do not satisfy these additional conditions, Lovejoy shows bounds on the value loss of these critical fractile policies versus an optimal policy. Lovejoy examines an N-horizon, discounted problem with zero lead time and linear costs. Excess demand is backordered although a certain fraction of backorders become observed lost sales. He specifically addresses exponential smoothing and Bayesian cases along with numerical examples. The cost is: $J = \sum_{t=1}^T \alpha^{t-1} C_t(x, a, w_t) + \alpha^T \sigma x_{T+1}$ where σ is the salvage value and a is the action that takes the inventory level to x . This problem can be solved by dynamic programming with the state space consisting of the current inventory and the sufficient statistic. The additional assumptions are

1. For $t = 1, 2, \dots, T$, the current cost function $C_t(x, a, w) = L_t x + K_t(a, w)$ for some constant L_t and function K_t , and with $L_{T+1} = \sigma$.
2. The transition equation for the inventory level is independent of the current stock level, *i.e.*, $x_{t+1} = g_t(x, a, w) = g_t(a, w)$.

If these assumptions hold in addition to Azoury's conditions then,

$G_t(y) = \int_0^\infty K_t(y, \xi) + \alpha L_{t+1} g_t(y, \xi) dF_t(\xi)$, $t = 1, 2, \dots, T$. Let y_t^* be the global minimizer of G_t which can be directly calculated from the problem data. The inventory problem can be solved in a series of T one-step minimizations of G_t . The optimal action at time t is $\delta_t^*(x_t, S_t) = q_t(S_t)y_t^*$. If excess inventory can be disposed of at cost, then Lovejoy's conditions are satisfied and the optimal policy is the solution of

$$F_t(y_t^*) = \pi = \frac{p - c(1 - \alpha\phi)}{h + p + \alpha(\phi - \gamma)} \quad (2.14)$$

where γ inventory deteriorates each period and ϕ backorders are lost each period. Lovejoy considers this same policy when disposal is not allowed, and it only meets those assumptions proposed by Azoury. He then shows the upper bound on the loss relative to optimal for using this simpler policy. He shows when it is feasible to hold onto stock rather than disposing it and the value of using a nondisposal policy.

Lovejoy demonstrates these models with examples using either exponential smoothing or Bayesian updating and makes several qualitative conclusions. For example, myopic policies perform better when perishability is high. As expected, myopic policies perform better as the precision of the information on the unknown parameter increases. The conditions of Lovejoy are similar to those we will see later in Lariviere and Porteus [21] because they assume stock deteriorates at the end of each cycle. They are both, in effect, eliminating the incoming stock level from the state space.

2.3.2 Censored Observations

The previous papers all assumed that the demand is always completely observed. However, when stockouts are lost, the demand is often only partially observed. That is, if there is no inventory left at the end of a period, the decision-maker may not know the amount of lost sales, in which case the demand process is only partially observed through the sales. This situation generates additional uncertainty about the demand process and probably occurs frequently in practice. It also provides an additional incentive to hold stock in order to gather more accurate demand information.

Although they do not study an adaptive inventory control model, Agrawal and Smith [1] and Nahmias [30] consider the problem with unobserved lost sales and make some observations that are beneficial to our problem. Agrawal and Smith argue that the negative binomial family often fits actual demand data well when the variance is high as in the case with many retailing cases. They describe how to fit a negative binomial distribution to unobserved lost sales data and demonstrate the goodness-of-fit for some actual sales data. Their technique, although not optimal for the Bayesian problem, could be used to update parameter estimates. Nahmias [30] also describes a parameter estimation procedure for unobserved lost sales using the normal distribution family. He develops a simple estimation procedure and demonstrates its accuracy compared to the maximum likelihood estimates. Again, managers should consider these procedures to update parameter estimates when there are censored observations.

Lariviere and Porteus [21] describe an optimal solution to a finite horizon problem that uses Bayesian techniques for the unobserved lost sales problem. This paper directly extends Azoury's results to the case where demand observations are censored due to unobserved lost sales. All costs and revenue are linear and lead time is zero. The identification of conjugate distributions is more difficult in this case. The authors model both a single retailer as well as a multiple market problem. The retailer's objective is to maximize his profits over a finite horizon. They assume that the retailer's excess inventory each period is lost. For the retailer, this decouples inventory levels between periods, but the state of information from the previous periods still couples the periods. The perishability assumption simplifies the problem significantly, but also limits application of the results. The retailers use a base stock policy that depends on the current information. The retailer can learn more about the true demand process by holding stocks at a higher level initially to more quickly establish a better estimate of the true demand. A true measure of demand occurs only when stock is left over at the end of the period. Braden and Freimer [6] define a family of "newsvendor" distributions that have a *fixed dimension* statistic even with *censored* observations. These distributions are of the form $\Phi(w; \tau) = 1 - e^{\eta(\tau)d(w)}$. The Weibull distribution $\{\phi(w; k, \tau) = \frac{k}{\tau}(\frac{w}{\tau})^{k-1}e^{-(\frac{w}{\tau})^k}; w > 0\}$ is the only "newsvendor" distribution that fits Azoury's requirements for a reduction in the state space and is therefore the distribution used by the authors (k known and τ unknown). The appropriate prior for τ at time t in this case is $\Gamma(a_t, b_t)$. The parameters are updated each period by the number of censored, m_c , and uncensored/exact, m_e ,

observations along with the last sales observation, w_{t-1} . Both parameters are a measure of market size, but only the shape parameter is a measure of the precision of information on the true demand distribution since the coefficient of variation for the gamma prior is $\sqrt{1/a_t}$. The parameters at time t are

$$a_t = a_1 + m_e \quad b_t = b_1 + \sum_{i=1}^t d(w_i) \quad \text{where} \quad d(w_i) = w_i^k \text{ for a fixed } k.$$

Notice that the shape parameter a_t only increases with exact observations. This fact is vital to the development of the policies. The sufficient statistic, S_t , is the vector of current parameters, (a_t, b_t) , for the prior. They show that, as in Azoury, the scale parameter can be eliminated from the analysis such that $J_t(a_t, b_t) = q(b_t)V_t(a_t)$. The simplified, single dimensional system comes from normalizing the initial inventory level and the expected demand. As in Azoury, the optimal order-up-to level is $\bar{x}_t = q(b_t)\gamma_t(a_t)$, where $\gamma_t(a_t)$ solves the reduced state system.

Lariviere and Porteus demonstrate several interesting results for managers to consider. If there is no penalty cost, the retailer will always prefer stockouts. When stockout penalties exist, the retailer will usually prefer excess inventory to reduce the level of uncertainty. Therefore, retailers should often stock more inventory than is necessary to maximize *current* profits in order to obtain more precise information on demand. For a given purchase cost, the optimal discounted expected return per unit of mean demand is strictly increasing in the number of exact observations. When stockout penalties are

small, a product that should be stocked will never be dropped despite poor sales. Moreover, although not necessarily intuitive, a product that is successful and stocks out might be dropped as a result. This is because the inability to increase the precision of information may lead to a situation where the risk of carrying too much inventory at a point closer to the end of the horizon is too high to stock the product.

Lariviere and Porteus extend the model to a multiple market problem by assuming there are j “independent” markets, the largest of which is noted as the primary market. The demand distributions are all Weibull, and the primary market has an unknown parameter with a gamma prior. The demand in the other markets are proportional to the demand in the primary market. Products cannot be dropped and the markets differ only in size. The decision maker has at least three stocking options. The “Naive” option treats all markets completely independently. A “Pooled Information” approach uses knowledge from all the markets to update the unknown parameter but the stocking decisions remain independent. The third “Joint Optimization” approach considers all interactions between markets in setting stocking levels. There are some interesting interactions. When the shape parameter is very low or high, there is a complementary effect that causes the overall stocking levels to be higher than when levels are considered independently. This is because each market has an incentive to stock higher levels to provide more information for other markets. On the other hand, in many (but not all) problems with intermediate values for the shape parameter, a substitution effect causes lower stocking levels in the joint problem. Since a market receives information from other

markets it has an incentive to stock at lower levels. Returns are higher when information is pooled as opposed to naive policies. If the shape parameter indicates any moderate or higher level of precision, then the joint optimization only performs nearly the same as the pooled information case. The authors also provide a persuasive argument that experimenting to gain information is better done in a smaller market than a larger market. The reasoning is that stocking one additional unit in a smaller market is more likely to lead to an exact observation than in the larger market. Due to the proportionality of demand, stocking one unit in a smaller market is equivalent to stocking one times the proportionality constant in the larger market. Finally, observations in a smaller market are at least as important as in the larger market.

2.4 Suboptimal Control

2.4.1 Continuous Parameter

Often times managers must make a series of decisions before any demand observations are available to update unknown parameters. Kaplan [16] considers a spare parts problem for a new system. In this case, it is often necessary to make some purchase decisions before fielding the new system, and thus, before actual demand is observed. In fact, several orders may be placed before fielding if the lead time is long, and once demands occur, the information is used to adjust inventory levels. Kaplan also assumes a fixed order cost and a periodic (s, S) control policy. He does not state restrictions on the holding and penalty costs, but we assume the results hold if they are convex. There is full

backordering and known constant lead times for procurement. He assumes the demand distribution is known and has an unknown parameter with an *a priori* distribution. He states that practicality limits choice of the prior to one from a conjugate family. An important issue for this problem is to determine how prefielding decisions change given that demand estimates change once fielding takes place. His dynamic programming formulation is the same as Equation (2.4) where $S_t = \sum_{i=1}^{t-1} w_i$.

$L(y|S_t)$ is the expected present value of backorder and holding costs in period $(t + L + 1)$, since inventory ordered at time t does not arrive until $t + L$. Kaplan assumes that the Bayesian updates will cease at some designated time and uses a heuristic to estimate the cost from that time until the end of the horizon. As long as this time is chosen carefully, it should not adversely affect the solution. Kaplan uses an (s, S) policy, although he does not prove its optimality. See Appendix for a sketch of this proof. He computes an (s, S) policy using the following steps. Let $y^*(x, S_t)$ be the optimal order-up-to level if the inventory level is x with sufficient statistic S_t . This is the value which minimizes Equation (2.4).

1. Find $y^*(0, S_t)$ and set $S = y^*(0, S_t)$
2. Find smallest x such that $y^*(x, S_t) = x$ and set $s = x - 1$.

Kaplan runs some experiments for Poisson demand with a gamma prior. A key part of his analysis was to examine how uncertainty about the unknown parameter affected the results. His results indicate that it is typically better to order less and wait for more information after the item has been fielded. The exact effect of uncertainty depends

on backorder costs, degree of uncertainty, and the time until fielding. There exists a natural tension between having lower reorder points until more information becomes available and setting higher reorder points to protect against shortages. Kaplan provides a heuristic which ignores the uncertainty in the prior that can be used as a prefielding policy. He then evaluates the increase in the lifetime costs of using this heuristic at two levels of uncertainty. Although this heuristic ignores the uncertainty in the system, it proved useful when backorder costs were relatively low.

Karmarkar [18] argues that inventory problems are often concerned with the determination of a fractile of the demand distribution, that is, the service level. Therefore, it is often sufficient to develop an accurate approximation of the right tail of the demand distribution and to disregard the rest of the distribution. Karmarkar provides an approach to approximate the tail distribution of demand. This procedure may be more effective than estimating the mean of a distribution and then inferring information about the tail from a hypothesized distribution. It may be more accurate to estimate the tail of the distribution from the beginning, especially when the distribution is difficult to approximate, e.g., bimodal distributions. His approach uses Bayesian analysis as well as heuristic smoothing methods. He demonstrates that his technique outperforms the Normal approximation for certain distributions, is simple, and uses intuitive statistics, although the method has some inherent bias. The approach assumes that the overall demand is a mixture of 2 probability density functions as shown in Figure (2.2).

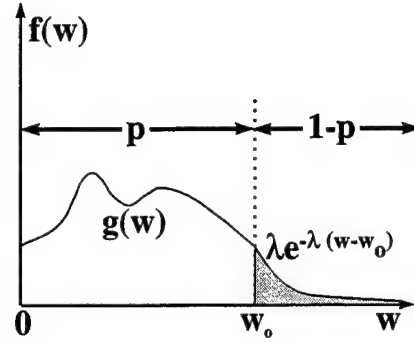


Figure 2.2: Mixture Model

Here, the demand density function is:

$$f(w) = \begin{cases} pg(w) & 0 \leq w \leq w_o \\ (1-p)\lambda e^{-\lambda(w-w_o)} & w > w_o \end{cases} \quad (2.15)$$

where $\int_0^{w_o} g(w)dw = 1$. Note that p is the probability that w is $\leq w_o$. The distribution to the left of w_o is of no concern in determining the stock level. Karmarkar suggests that w_o should be near the 80% fractile. Above the set point w_o , the tail distribution is a translated exponential which accurately approximates the tail of many distributions. Given that S_L is the desired service level, with known parameters w_o , λ , and p , the order-up-to level is

$$\bar{x} = w_o + \frac{1}{\lambda} \ln \left[\frac{1-p}{1-S_L} \right] \quad (2.16)$$

Since p and λ are not generally known, Karmarkar provides two estimates of this order-up-to level. The first he calls an exact estimator which is derived from a Bayesian analysis based on the distribution of demand, *unconditional* on p and λ . It accounts for the extra uncertainty that is present in the system due to estimating p and λ . The exact estimator is

$$\bar{x}_1 = w_o + w_g \left[\left(\frac{1 - \bar{p}}{1 - S_L} \right)^{1/n_g} - 1 \right] \quad \text{where} \quad \bar{p} = \frac{n_l}{t}, \quad (2.17)$$

there are t total observations, n_l is the number of points less than or equal to w_o , n_g is the number of points greater than w_o , and $w_g = E[w|w > w_o]$. The derivation of the exact estimator assumes a prior distribution on both p (beta with parameters w_g and n_g) and λ (gamma with parameters n_l and t). An approximate estimator is given by estimating \bar{p} and $\bar{\lambda}$ and using Equation (2.16). This approach is a type of certainty equivalent control that assumes \bar{p} and $\bar{\lambda}$ are exact. Karmarkar tests these estimators against a standard Normal distribution using both MAD and variance for four distributions (Normal, exponential, Erlang, and Normal with zero spike). The mixture estimates either outperform the standard approaches or perform satisfactorily for these distributions. The approximate estimator always provides a critical fractile estimate that is below the exact estimate from Equation (2.17).

2.4.2 Discrete Parameter

Lovejoy [28] models the problem with a discrete number of possible demand distributions. He computes bounds on the optimal policy, and he shows how to iteratively tighten both the upper and lower bounds on the optimal policy. He also shows how to generate a suboptimal policy in the process of computing these bounds. His model has an infinite horizon and allows fixed ordering costs. Holding and shortage costs are linear and lead time is zero. Excess demand is lost, and actual demand may or may not be observed exactly.

The problem is modeled as a composite-state Partially Observed Markov Decision Process (POMDP). Some states of the system are known with certainty while others are only partially observed. That is, the inventory level is known with certainty, but the demand distribution in effect, $\tau \in \{1, \dots, n\}$, is unknown. The objective is to maximize the expected discounted reward over the infinite horizon. The unknown parameter, τ , is the demand distribution that is in effect. The prior for τ is represented by π , and define the set $\Pi = \{\pi \in R^n : \pi_i \geq 0, \sum_{i=1}^n \pi_i = 1\}$. Lovejoy uses properties of POMDPs developed by Smallwood and Sondik [35] and extends suboptimal techniques used by Van Hee [41] to suboptimally solve this POMDP problem. If a partially observable state becomes known with certainty, the problem is then *fully observed* and assumed solvable by conventional dynamic procedures. Van Hee exploits this fact to show how upper and lower bounds can be developed by solving *fully observed* problems.

Lovejoy develops parallel procedures that are more powerful because they use approximate value functions from “action invariant sets” to set bounds and adjust them iteratively. The use of these sets gives a larger class of cost functions to approximate the optimal policy. Because it is well known that the maximum reward over the infinite horizon will come from a stationary Markov policy, let $J(x, \pi)$ represent the total discounted expected reward function for an inventory level x and prior π , that is $J(x, \pi) = E[\sum_{t=1}^{\infty} \alpha^{t-1} r(x_t, \tau, \delta(I_t))]$ where the true demand distribution is τ , the action is $\delta(I_t)$, the discount rate α , and $r(\cdot)$ is the reward. Define the value function $C(x, \pi, a, J) = \sum_{i=1}^n \pi_i r(x, \tau, a) + E\{\alpha J(x_{t+1}, \pi_{t+1})\}$. Lovejoy defines a mapping $H_\delta J(x, \pi) = C(x, \pi, \delta(x, \pi), J)$ as the expected current reward plus the future reward using value function J and a feasible policy, δ . Finally, let $HJ(x, \pi)$ be the maximum present and future reward over all possible actions, a , at the current stage and let J^* be the optimal value function. This implies that a specific policy δ is optimal if and only if $H_\delta J^* = HJ^*$. Although it is impractical to do so since the state space in Π is countably infinite, the problem could theoretically be solved by dynamic programming. The following key relationship is used to generate bounds and can be easily verified (let a caret $\hat{\cdot}$ denote a fully observed problem and e_i is a unit vector with a 1 in the i th position)

$$J_\delta(x, \pi) = \sum_{i=1}^n \pi_i J_\delta(x, e_i) = \sum_{i=1}^n \pi_i \hat{J}_{i, \delta}(x). \quad (2.18)$$

Van Hee showed that a lower bound, $J_l(x, \pi)$, can be determined by solving the n fully observed problems. Let $\gamma_\delta(x) = (J_\delta(x, e_i))_{i=1}^n$ be an n -vector representing the rewards of policy δ and value function J . Each component of the vector represents the reward if the true demand distribution is i . Then, $J_l(x, \pi) = \max_{\gamma \in \Gamma(x)} \langle \gamma \cdot \pi \rangle$. Van Hee shows that a one-step dynamic recursion using J_l results in reward HJ_l and the following bounds hold: $J_l \leq HJ_l \leq J^*$. The upper bound is also found using by solving the n fully observed problems where $J_u(x, \pi) = \sum_{i=1}^n \pi_i J^*(x, e_i)$. Again, a one step recursion gives a reward of HJ_u and the following bounds are obtained:

$$J_l \leq HJ_l \leq J^* \leq HJ_u \leq J_u \quad (2.19)$$

These bounds can be iteratively tightened in theory by additional dynamic programming recursions since there are a finite number of states. However, the state space grows exponentially making the approach practically intractable.

Lovejoy first extends the above results to show that a collection of invariant sets ($\Psi(x)$) can be used to generate approximate value functions. These functions can be calculated at any point over the entire state space ($X \times \Pi$) rather than at the extreme points, e_i . A collection of sets is invariant if the associated policies (actions) are a function only of the observed state. The functions that are generated from these invariant sets preserve the bounds shown in Equation (2.19). Given a collection of invariant sets, $\Psi(x)$, a lower bound $J_l = \max_{\psi \in \Psi(x)} \langle \psi \cdot \pi \rangle$ where the n -vector $\psi \in \Psi(x)$. The lower bound is

found analogously to Van Hee, although now the collection of invariant sets may include vectors, ψ , generated at points other than the extreme points, $\{e_i\}$. The upper bound is found in the same manner as in Van Hee ($J_u(x, \pi) = \sum_{i=1}^n \pi_i V^*(x, e_i)$). Once again, a single dynamic recursion using J_l and J_u can improve these bounds. Bounds are tight near the extreme points as well as at points where observations are highly informative about the underlying demand distribution. They may also be tight when rewards are not observed such as in the case of some lost sales problems when only sales, not demand, is observed. Since this technique should be used only when the partially observed states have a significant effect on the rewards and transitions, a single dynamic programming recursion (observation) should make a significant impact on the value functions.

Lovejoy demonstrates his technique with an inventory example. The optimal policy has an (s, S) structure that is a function of the current prior, *i.e.*, $(s(\pi), S(\pi))$. The example demonstrates that stocking levels will generally be higher when only sales are observed because inventory has inherent informational value. Thus, the example illustrates the concept of probing, that is, stocking more inventory to get a more precise estimate of demand when sales, not demand, are observed. He also shows when managers will prefer not to order (policy is $(0, 0)$) unless more information is received.

Lovejoy then presents a finite grid approximation method to tighten the bounds and determine suboptimal policies. First, for an invariant collection of sets, the function $J_l(x, \pi)$ is piecewise linear and convex over Π . This property, as Sondik and Smallwood [35] discuss, allows new vectors to be added to the set in order to improve the

lower bound. Lovejoy suggests using any finite set of points in Π . If the new value functions are denoted by J' and HJ' , then $J_l \leq J' \leq HJ' \leq J^*$. The discrete points chosen in Π should be points where there is a large difference between the upper and lower bounds. At each dynamic programming recursion a suboptimal policy is generated to improve the lower bound, and the upper bounds can be tightened by choosing points that form a triangulation of Π and computing $H_u J(x, \pi) = \sum_{i=1}^n \lambda_i HJ(x, v_i)$ where v_i represents the selected points and λ_i are such that $\pi = \sum \lambda_i v_i$. Then, the following results hold: $H[H_u J_u] \leq HJ_u \leq J_u$ and if $J'' = H_u J_u$, then $J^* \leq HJ'' \leq J''$.

Lovejoy extends these results to cases in which some parameters are nonstationary. The underlying demand distribution remains stationary and the horizon length is finite. Again he shows how to calculate suboptimal policies and then iteratively adjust both the upper and lower bounds. The procedures are very similar to those for the stationary case, except that the procedure must account for changes in parameter values over time.

2.5 Conclusion

Managers can maximize profits by very judicious control of their inventory. To do so, they do not need to automatically assume a system of certainty equivalence control. This task has become increasingly difficult in many of today's volatile markets which are marked by high levels of uncertainty. Despite the inherent difficulties, there are practical tools and theory available to eliminate some of the costly effects of uncertainty. This chapter has reviewed several optimal control policies for both uncensored and censored

demand observations and has presented a summary of key properties for the solutions to these problems. For some problems, state space reductions make partial information problems as solvable as their perfect information counterparts. In most practical cases, however, problems must be solved suboptimally.

An analysis of the current research makes it evident that there is room for much more research in this area. Perhaps the largest area of potential research is the extension of the current theory and models to “nonstationary” demand problems in which the underlying demand distribution changes over time. Many actual inventory control problems are characterized by partial information plus nonstationary demand. This extension, combined with partial information, complicates the calculation of an optimal or suboptimal solution. However, with the advances in computer memory and speed, solutions to these problems should be possible with well designed algorithms.

Although there are potentially significant financial benefits to adaptive control procedures, several fundamental obstacles prevent wide use of these techniques. First, managers lack a set of practical tools that they can use to quantify the uncertainty and to adaptively control inventory in the face of this uncertainty. Managers must be provided these tools as well as an understanding of the effect of the assumptions that they must make. For example, managers must know how to determine the *a priori* beliefs for the unknown parameters. While doing so, they need to know how this assumption affects the solution they obtain, *i.e.*, the Bayesian robustness of the prior. Second, managers often make parameter estimates and then rely on some form of certainty equivalence

control (CEC). Managers must understand the expected level of error when using such a procedure. For stationary problems, the adaptive control procedures will converge to a CEC solution. Thus, one must know when it is economically prudent to switch from an adaptive control procedure to a CEC. Once a CEC is used, a valid procedure is needed to identify when use of CEC is no longer economically sound, such as in the tail of a product life cycle. Effective tools are needed to determine when such changes in strategy should be made. Third, managers may choose to forego adaptive procedures due to the restrictions on the prior and demand distributions. For the continuous parameter cases, computational resources normally require that the demand and prior distributions come from conjugate families. Managers should know when their demand process can be accurately approximated by these families and the cost of assuming that they do.

Finally, in some cases there may not be a conjugate family that adequately approximates the real problem. In these cases, a bounded discretized prior can be developed for any desired level of accuracy along with a set of bounded, discretized demand distributions. Using a discrete model and solution strategy similar to that of Lovejoy [28], the posterior distribution can now be easily calculated as well as a suboptimal control policy. Thus, this discretization approach constructs a policy using suboptimal control for accurate prior and demand distributions whereas the continuous modeling approach applies optimal control to, in these cases, less accurate prior and demand distributions. An interesting area of research would be to determine when the discretization approach

provides better control policies than the continuous model when the necessary distribution assumptions are violated.

We have reviewed the key literature for this adaptive inventory control problem and identified some of its salient characteristics and areas for further research. The problem area is important for a large number of actual inventory management environments, and this number continues to grow. In conclusion, it is clear that there are many interesting and relevant areas of research for this fundamental inventory problem, and we hope that this review will stimulate that research.

2.A Optimality of (s, S) policy

The inventory problem with fixed purchase costs, linear holding and penalty costs, and perfect information has an optimal (s, S) policy. Although Kaplan [16] uses an (s, S) policy for his problem, he does not show its optimality. Lovejoy [28] suggests that the proof of the optimality of the (s, S) policy for the partial information problem is an extension of the perfect information proof. Bertsekas [5] provides a proof for the *perfect* information problem, and we point out the essential parts of the extension of this proof to the *partial* information problem. As shown in Figure (2.3), the cost function is not convex, so the proof strategy is to show that for all t and π_t , $J_t(x_t; \pi_t)$ is *K-convex* in x_t . If a continuous function g is K-convex and $g(y) \rightarrow \infty$ as $|y| \rightarrow \infty$ then by Lemma 2.1 in Bertsekas (page 149), an (s, S) policy is optimal. That is, K-convexity proves the uniqueness of both s and S . Otherwise, the situation in Figure (2.4) could occur which would imply multiple intervals for reordering.

We present the following definition and lemma from Bertsekas for clarity.

Definition. A real function g is *K-convex* if $\forall z \geq 0, b > 0, K \geq 0, y$; then $K + g(z + y) \geq g(y) + z(\frac{g(y) - g(y-b)}{b})$.

Lemma 2.1.

- (a). A real-valued convex function g is also 0-convex and K-convex for all $K \geq 0$.
- (b). If $g_1(y)$ and $g_2(y)$ are K-convex and L-convex, then $\alpha g_1(y) + \beta g_2(y)$ is $\alpha K + \beta L$ - convex for all $\alpha, \beta > 0$.
- (c). If $g(y)$ is K-convex and w is a random variable, then $E_w\{g(y - w)\}$ is also K-convex if $E_w\{|g(y - w)|\} < \infty$ for all y .
- (d). If g is a continuous K-convex function and $g(y) \rightarrow \infty$ as $|y| \rightarrow \infty$, then there exist scalars s and S with $s < S$ such that

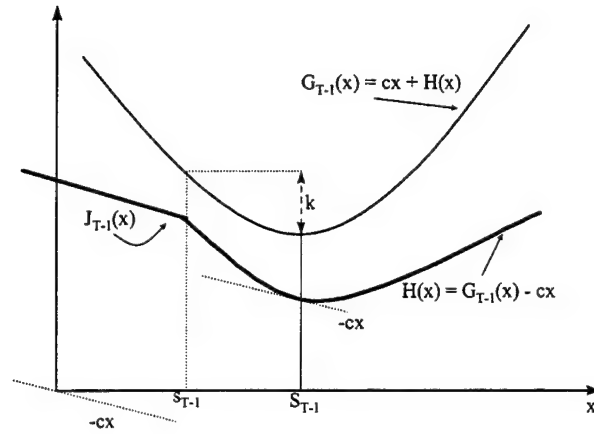


Figure 2.3: Cost Function with Fixed Order Cost

- (1) $g(S) \leq g(y)$, for all y ;
- (2) $g(S) + K = g(s) < G(y)$, for all $y < S$;
- (3) $g(y)$ is decreasing on $(-\infty, s)$;
- (4) $g(y) \leq g(z) + K$ for all y, z with $s \leq y \leq z$.

Let

$$G_t(y; \pi_t) = cy + H(y; \pi_t) + E_w\{J_{t+1}(y - w; \pi_{t+1})\} \quad (2.20)$$

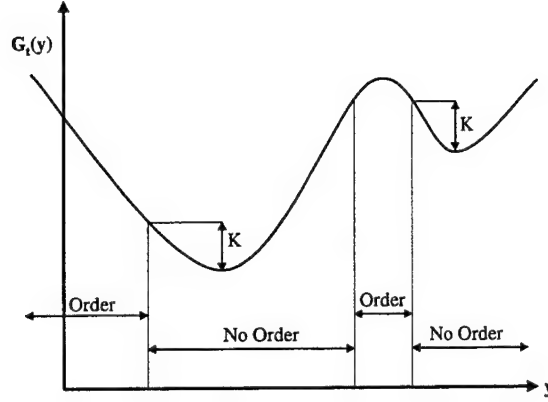


Figure 2.4: Possible Cost Function with Fixed Order Cost

where $H(y; \pi_t)$ is the expected holding and penalty costs in period t . If $J_T(x_t; \pi_T) = 0$, then for fixed π , G_{T-1} is clearly convex. The cost function is

$$J_{T-1}(x_{T-1}; \pi_{T-1}) = \begin{cases} K + G_{T-1}(S_{T-1}; \pi_{T-1}) - cx & \text{if } x < s_{T-1}; \\ G_{T-1}(x; \pi_{T-1}) - cx & \text{if } x \geq s_{T-1}. \end{cases} \quad (2.21)$$

Given that G_{T-1} is K-convex, to show that Equation (2.21) is K-convex, we must show that $\forall z \geq 0, b > 0, K \geq 0, y$,

$$K + J_{T-1}(y + z; \pi_{T-1}) \geq J_{T-1}(y; \pi_{T-1}) + z \left(\frac{J_{T-1}(y; \pi_{T-1}) - J_{T-1}(y - b; \pi_{T-1})}{b} \right). \quad (2.22)$$

There are three cases which must be examined. The first case is for $y \geq s_{T-1}$. If $y - b \geq s_{T-1}$, then $J_{T-1}(x_{T-1}; \pi_{T-1})$ is clearly K-convex since it is the sum of a K-convex function and a linear function by Equation (2.21) and Lemma 2.1(b). If $y - b < s_{T-1}$, Equation (2.22) can be written as

$$K + G_{T-1}(y + z; \pi_{T-1}) \geq G_{T-1}(y; \pi_{T-1}) + z \left(\frac{G_{T-1}(y; \pi_{T-1}) - G_{T-1}(s_{T-1}; \pi_{T-1})}{b} \right). \quad (2.23)$$

Then y is such that $G_{T-1}(y; \pi_{T-1})$ is either \geq or $< G_{T-1}(s_{T-1}; \pi_{T-1})$. In either case, Equations (2.22) and (2.23) hold, and $J_{T-1}(x_{T-1}; \pi_{T-1})$ is K-convex. The second case is for $y \leq y + z < s_{T-1}$. $J_{T-1}(x_{T-1}; \pi_{T-1})$ is linear here by Equation (2.21) and therefore K-convex. The third case is $y < s_{T-1} < y + z$. Again, Equation (2.22) holds and $J_{T-1}(x_{T-1}; \pi_{T-1})$ is K-convex. We have shown that due to the K-convexity of G_{T-1} that $J_{T-1}(x_{T-1}; \pi_{T-1})$ is also K-convex as well as continuous. Extending Lemma 2.1(c), if $g(y; \pi)$ is K-convex for fixed π and w is a random variable, then $E_w\{g(y - w); T(\pi, w)\}$ is also K-convex when $E_w\{|g(y - w)|; T(\pi, w)\} < \infty$. Assuming that demand is bounded, then the expectation over π preserves the K-convexity. Therefore, G_{T-2} is also K-convex, continuous, and $G_{T-2}(y; \pi_{T-2}) \rightarrow \infty$ as $|y| \rightarrow \infty$. We can show that for all t that $G_t(y; \pi_t)$ is K-convex, continuous, and $G_t(y; \pi_t) \rightarrow \infty$ as $|y| \rightarrow \infty$ which implies that $J_t(x_t; \pi_t)$ is also K-convex. By Lemma 2.1(d), an (s, S) policy is optimal. \square

CHAPTER 3

POLICIES FOR NON-STATIONARY DEMAND AND PARTIAL INFORMATION

3.1 Introduction

Product demand is often characterized by partially observed, non-stationary demand. The demand process is partially observed because the distribution of demand in a given period is not known with certainty, and it is nonstationary because the distribution may randomly change over time. When these conditions exist, managers must estimate the unknown demand distribution, and often some type of certainty equivalent control (CEC) that assumes the estimate is exact is used to solve the problem. We present compelling research that demonstrates that there exist other practical suboptimal control policies to solve realistic instances of this problem without assuming such an estimate. Furthermore, these control policies achieve much better performance in some cases than the CEC policies found in practice. These suboptimal procedures include Open-Loop Feedback Control and Limited Look-Ahead Control. Both of these policies account for more of the inherent uncertainty in the demand process and, therefore, often outperform the CEC policies. We examine a finite horizon problem with positive lead time, full backordering, and linear holding and backorder costs. We model the demand process as a composite-state, partially observed Markov decision process. The demand distribution is assumed to belong to a known, discrete set of candidate distributions. A

known, discrete *a priori* distribution is assumed for this set that is used along with subsequent observations to compute a *posterior* distribution. We test these control policies over a wide range of problem instances and compare them with an optimal policy and three CEC policies.

3.2 Literature Review

The inventory control literature can be categorized into four classes. First, there are problems with complete state information and stationary demand which form the body of most of the classical inventory theory. Lee and Nahmias [23] summarize this classical theory quite well. Secondly, there are problems with complete information and non-stationary demand. These problems, although they are more complex, have been studied fairly well also. For example, Anupindi, Morton, and Pentico [2] provide four heuristics for a nonstationary stochastic lead time problem. Much less research is available for the other two categories of problems, those with partial information and stationary or non-stationary demand. This chapter will focus on the case of non-stationary, stochastic demand with incomplete information. But first, we review some of the literature for related problems.

One of the earliest cases of a stochastic non-stationary demand problem is found in Hadley and Whitin [13]. They formulate a solution to a final inventory problem for which there is a known obsolescence date and Poisson demand. Shortly prior to Hadley and Whitin, Karlin [17] analyzes a dynamic system (non-stationary and stochastic) in

which the distribution of demand can vary each period, and it is an extension of the Arrow-Harris-Marshak [3] dynamic inventory model. Karlin's main contribution is to qualitatively show how the critical number (optimal inventory level) varies over time as a function of the demand densities. He determined that when the demand densities decrease stochastically in consecutive periods, the critical number decreases. If the densities increase, the critical number may or may not increase. Finally, if the densities increase in successive periods while the critical number decreases, then the critical number decreases further in the following period. In his final section, he discusses cases of "partial information" in which an unknown parameter of the demand distribution has an *a priori* distribution and the distribution is stationary. He analyzes some cases in which a single sufficient statistic exists, namely the exponential family (including gamma, Poisson, and negative binomial) and the range family. Karlin then presents four theorems in which he shows characteristics of optimal policies for both the exponential and range families. Veinott [42] extended Karlin [17] by showing that when one density is a translate of another, it is possible to relax the requirement for stochastic ordering to show the relationship between two sets of critical numbers. He also shows how to change a zero lead time problem with no backlogging to one with backlogging.

More recently, Song and Zipkin [37] discuss a general modeling framework in a fluctuating demand environment. They discuss characteristics of optimal policies and describe two different algorithms to solve the linear cost problem. Their paper showed that optimal policies depend on the unit cost, holding cost, and penalty cost as in

the newsvendor problem. Later, Song and Zipkin [38] demonstrate how inventories should be managed in the face of possible obsolescence. They assume that a continuous time, non-increasing discrete state Markov chain describes the demand process. One critical assumption they make is that the *current demand state is always known exactly*. They calculate and compare an optimal policy against two heuristics. One heuristic, a blind policy, forecasts demand over a lead time assuming that there is no change in demand over the lead time. The second heuristic, a limited look-ahead policy, manages demand over a lead time and accounts for demand changes over the lead time but not beyond. Stephen Graves [12] proposes an adaptive base-stock inventory policy for a non-stationary problem. His demand model is an integrated moving average model of order (0,1,1): $d_1 = \mu + \epsilon_1$ and $d_{t+1} = d_t - (1 - \alpha)\epsilon_t + \epsilon_{t+1}$ for $t = 1, 2, \dots$. He characterizes results for his assumed policy without asserting its optimality, and he uses an exponential smoothing procedure to estimate the demand. This problem has complete information because the distribution of the demand in a period is fully determined by the ARIMA model and the observed demand in the previous period.

Problems with partial information are significantly more difficult to solve. Stationary, partial information problems may be solved by Bayesian methods or other adaptive procedures. Herbert Scarf [33] was a pioneer in the use of Bayesian techniques for inventory control. He described an inventory problem with linear holding and penalty costs. The demand density was described by a density $\phi(\xi, \omega)$ with unknown parameter ω . The

parameter ω , however, has a known *a priori* distribution. The demand density was restricted to the exponential family because a single sufficient statistic, S , can summarize all prior demand information. For the exponential family, the sufficient statistic, S , is the average demand over the previous periods. Scarf shows that the optimal stock level can be calculated recursively with knowledge of the current inventory level, x , and the sufficient statistic, S . In a subsequent paper [34], he shows that in some instances, the critical level can be determined as a function of one variable if the demand distribution is in the gamma family. Katy Azoury [4] demonstrated that other problems can also be solved with a one-dimensional state space when there is a natural conjugate family. If the demand distribution has a fixed dimension sufficient statistic, and the distribution of the unknown parameter comes from a conjugate family, then the posterior distribution for the unknown parameter is also in that same family. She then presents two conditions that show the model can be reduced to a one-dimensional state space and solved by a dynamic program.

Lovejoy [25] uses a myopic parameter adaptive technique and a simple inventory policy based upon a critical fractile. Specifically, he uses exponential smoothing and Bayesian updating of parameter estimates. He also determines bounds on the value loss relative to optimal costs when using his policy. Lovejoy [28] provides additional insights into the stationary, partial information problem. These insights deal with bounds for certain suboptimal policies, and he shows how it is possible to iteratively tighten these bounds. He also addresses a problem in which some parameters may be nonstationary,

although the underlying demand distribution remains stationary. Lovejoy introduces the concept of a composite-state partially observed Markov decision process in which the inventory level is observed completely while the demand distribution is observed partially.

Gallego, Ryan and Simchi-Levi [10] demonstrate a Min-Max technique for analysis of various distribution free finite horizon models for which the distribution is specified by a limited number of parameters such as the mean and variance or a set of percentiles of demand. They provide a polynomial algorithm to minimize the maximum expected costs over all possible demand distributions with the given parameters.

Lariviere and Porteus [20] consider Bayesian techniques for a lost sales problem in which sales, not true demand, are observed. Because of lost sales, the retailer can learn more about the true demand process by holding inventory at a higher level initially to quickly establish a better estimate of the true demand. The process of finding conjugate distributions is more difficult when there is censored data due to lost sales. Lariviere and Porteus show that a Weibull density for demand with a gamma prior on one of the unknown parameters provides the necessary conditions for a known posterior distribution on demand. Agrawal and Smith [1] also examine lost sales and contend that the negative binomial family provides a very good fit for actual retailing sales data. They describe procedures to estimate the unknown parameter of the negative binomial distribution. Nahmias [30] also proposes procedures for demand estimation in the case of lost sales assuming normally distributed demand. Karmarkar [18] shows that inventory problems

are often concerned with the determination of a fractile of the demand distribution which is equivalent to determining a service level and demonstrates one way to estimate the critical fractile. Pierskalla [31] studied the stochastic demand problem in which demand is stationary but with an unknown termination date. He shows that the distribution of the date of obsolescence can greatly affect the critical number. Finally, Kaplan [16] uses a Bayesian approach to update demand information for a procurement problem. First, an initial procurement is made before a new item is put into service. After the part enters service, demand information is used to provide a more accurate forecast.

Problems with non-stationary demand and partial information present an even more complex situation. Little direct work has been done in this area, but, there is great potential impact for research on this class of problems. One paper that considers a problem in this class is by Kurawarwala and Matsuo [19]. They present a growth model to estimate the parameters of a demand process over its entire life cycle. In their base case, production decisions are made at the beginning of the problem for the entire life cycle. They present a technique to initially estimate the parameters of their forecasting model. However, they do not thoroughly address the issue of revising these estimates using new observations.

We model the problem of non-stationary demand with partial information as a composite-state, partially observed Markov decision process (CPOMDP), see Section 3.3. Partially observed Markov decision processes have been well studied in the past, but with few applications in inventory and production control. However, it composes

a body of theory with some direct application to our problem. Monahan [29] presents a good survey of POMDP. Smallwood and Sondik [35] show that for a finite horizon POMDP maximization problem the objective function is piece-wise linear and convex in the current state probabilities. They then provide a dynamic programming algorithm that uses this property to solve the problem. Lovejoy [24] presents monotonicity results for certain POMDP problems. White and Sherer [43] present three algorithms for the POMDP that are faster than that of Smallwood and Sondik. However, the problem sizes are very limited in their study. White and Sherer [44] propose a heuristic in which only the M most recent observations and actions are used to make the current decision. They also introduce the concept of ergodicity and discuss how this factor indicates the potential value of historical information.

3.3 Optimal Control Model and Insights

We first state our basic model that will serve as a reference model for further discussion. We assume that the demand in any given period arises from one of a finite collection of N probability distributions. However, the decision maker may not have complete certainty about which of the distributions generates the demand in a given period. Furthermore, the distribution may randomly change from one period to the next. Any pattern of transitions from one distribution to another can be modeled. For this basic model, we assume that the planning horizon is finite, stockouts are always backordered, the order leadtime is either zero or a known positive constant, and there is no

fixed order cost. However, this model can be extended to allow a positive fixed order cost and observed or unobserved lost sales. If lost sales are observed, the stockout penalty cost will normally be adjusted accordingly. If the lost sales are unobserved, the revision of the prior for the unknown parameter is more complicated. Lost sales cases also require an augmented state space which can significantly complicate the computation of costs. We define in Table 3.1 the notation that will be used to describe the model.

N	=	number of distributions
M	=	maximum demand
S	=	stock level
h	=	linear holding cost,
p	=	linear stockout penalty cost,
c	=	linear procurement cost,
T	=	end of planning horizon,
l	=	lead time,
d_t	=	demand distribution in effect for period t ,
p_{ij}	=	$P[d_{t+1} = j d_t = i]$,
P	=	(p_{ij}) , the transition probability matrix for d_t ,
w_t	=	demand realization in period t ,
r_{jk}	=	$P[w_t = k d_t = j]$,
I_t	=	vector of information available up to period t ,
π_{it}	=	$P[d_t = i I_t]$,
π_t	=	(π_{it}) ,
x_t	=	inventory level at the beginning of period t ,
u_t	=	inventory position at the beginning of period t ,
a_t	=	order quantity in period t .
$\bar{a}_{t,l}$	=	vector of outstanding orders,

Table 3.1: Notation

Although we specify linear holding and stockout costs in this basic model, many of the results hold for convex costs. The demand state process $\{d_t\}$ is not known with certainty and is referred to as the *core process*. It is modeled as a finite state Markov

chain, $d_t \in \{1, \dots, N\}$, governed by the transition matrix P . If $d_t = j$, the probability distribution for demand in t is r_j , i.e., $P[w_t = k | d_t = j] = r_{jk}$ for $k = 0, 1, \dots, M$. Although we assume that the matrix P is stationary, the model could be readily extended to allow a non-stationary transition process. Note that there is a great deal of flexibility in modeling the core process. The case of a stationary demand process is modeled using $P = I$, the identity matrix, while an upper triangular matrix, P_u , or a lower triangular matrix, P_l , can be used to model a demand trend. A slight modification of the upper or lower triangular matrices can be used to model cyclical demand. The core process is partially observed through the *measurement process* $\{w_t\}$. However, in the case of unobservable lost sales we assume that the core process is observed through the sales rather than the actual demand; therefore, the measurement process is different. For the basic model, the vector of information available at time t is $I_t = (w_0, w_1, \dots, w_{t-1})$. Also known is π_t and x_t where π_0 is the initial prior distribution that is externally specified. In practice, π_0 will be based on previous experience with similar products or preliminary market analysis. Rhenius [32] has shown that the prior distribution π_t is a sufficient statistic for I_t . Therefore, the problem can be viewed as a Markov decision process on the state space (x_t, π_t) . The vector π_t characterizes the current belief of the distribution of d_t given all prior observations of the information process $\{w_t\}$. The system dynamics for our CPOMDP model are described in Figure 3.1. The control process begins with an information vector, I_t , which includes a prior distribution, π_t , for period t , and then selects an order quantity, a_t . For positive lead time, an order placed l periods in the

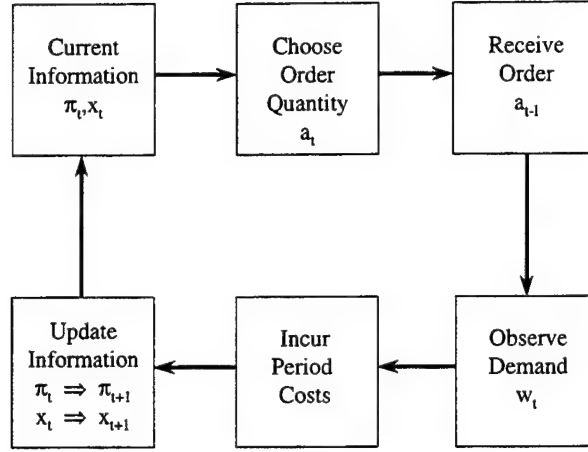


Figure 3.1: Control Process

past is received. The demand, w_t , then occurs, and the inventory costs for period t are incurred. The time index advances to $t + 1$, and the core process advances from d_t to d_{t+1} according to the transition matrix P . The new prior distribution π_{t+1} is computed using π_t, w_t and P . Note that our model differs from the typical POMDP formulation (Monahan [29], Bertsekas [5]) in that the most recent observation, w_t , used to update π_{t+1} is an observation of d_t and not an observation of d_{t+1} . Therefore, the transition equation for our model is

$$\pi_{i,t+1} = T_i(\pi_t | w_t = k) = \frac{\sum_{j=1}^N \pi_{jt} r_{jk} p_{ji}}{\sum_{j=1}^N \pi_{jt} r_{jk}}. \quad (3.1)$$

If we assume that all stockouts are backordered and that the leadtime is zero, the transition equation for the inventory level is

$$x_{t+1} = x_t + a_t - w_t. \quad (3.2)$$

For an inventory level v , define the single period expected inventory cost function as

$$G_t(v|\pi_t) = E_{w_t|\pi_t}[p \max\{0, w_t - v\} + h \max\{0, v - w_t\}]; \quad (3.3)$$

then the optimal stochastic dynamic programming cost-to-go function is

$$J_t(x_t, \pi_t) = \min_{a_t \geq 0} \{ca_t + G_t(x_t + a_t|\pi_t) + E_{w_t|\pi_t}[J_{t+1}(x_t + a_t - w_t, T(\pi_t|w_t))]\},$$

$$t = 0, \dots, T, \quad (3.4)$$

where $t = 0$ indexes the first period and the sequence of calculations is $T, T-1, \dots, 1, 0$.

The terminal cost function is defined as $J_{T+1}(x_{T+1}, \pi_{T+1}) = -cx_{T+1}$ so that there is no lost value for the terminal inventory. However, the results for this model apply if $J_{T+1}(x_{T+1}, \pi_{T+1})$ is a convex function of x_t and is a piece-wise linear concave function

of π_t . Alternatively, if we define $S_t = x_t + a_t$, the problem can be stated as

$$J_t(x_t, \pi_t) = -cx_t + \min_{S_t \geq x_t} \{cS_t + G_t(S_t|\pi_t) + E_{w_t|\pi_t}[J_{t+1}(S_t - w_t, T(\pi_t|w_t))]\},$$

$$t = 0, \dots, T. \quad (3.5)$$

Our model permits a positive and known, integer leadtime, $l > 0$. We assume that the delivery quantities are equal to the order quantities and are completely known. This extension requires that we augment the state space to include the outstanding orders at time t . Define the vector $\bar{a}_{t,l} = (a_{t-l}, a_{t-l+1}, \dots, a_{t-1})$; then the state vector at time t is $(x_t, \pi_t, \bar{a}_{t,l})$. The inventory transition equation is $x_{t+1} = x_t + a_{t-l} - w_t$, and the order transition equation is $\bar{a}_{t+1,l} = \bar{a}_{t,l} - \{a_{t-l}\} + \{a_t\}$. For an inventory level v , define the single period expected inventory cost function as

$$G_t(v|\pi_t, l) = E_{w_{t,l}|\pi_t}[p \max\{0, w_{t,l} - v\} + h \max\{0, v - w_{t,l}\}], \quad (3.6)$$

where $w_{t,l} = \sum_{n=0}^l w_{t+n}$. This cost function accounts for the positive leadtime in computing the expected inventory costs for the inventory level at time $t + l$. The non-stationary nature of the core process complicates the calculation of the conditional distribution $(w_{t,l}|\pi_t)$. The matrix P^n is the n -step transition probability matrix, and its transpose is denoted $(P^n)'$. For $n = 1, \dots, l$, the conditional distribution of π_{t+n} given π_t can be computed as $(\pi_{t+n}|\pi_t) = (P^n)'\pi_t$, and the conditional distribution of w_{t+n}

given π_t is $(w_{t+n}|\pi_t) = (\pi_{t+n}|\pi_t)'R$ where $R = (r_{jk})$. Finally, the conditional distribution is calculated for the convolution $(w_{t,l}|\pi_t) = \sum_{n=0}^l (w_{t+n}|\pi_t)$. The cost-to-go function becomes

$$J_t(x_t, \pi_t, \bar{a}_{t,l}) = \min_{a_t \geq 0} \{ca_t + G_t(x_t + a_{t-l}|\pi_t) + E_{w_t|\pi_t}[J_{t+1}(x_t + a_{t-l} - w_t, T(\pi_t|w_t), \bar{a}_{t+1,l})]\}, \quad t = 0, \dots, T. \quad (3.7)$$

This problem can be simplified by formulating it in terms of the *inventory position*, which is the net inventory level (on-hand less backorders) plus inventory on-order at the beginning of t , $u_t = x_t + \sum_{n=1}^l a_{t-n}$. The inventory position follows the transition equation $u_{t+1} = u_t + a_t - w_t$, and if we redefine $S_t = u_t + a_t$ then the DP recursion becomes

$$J_t(u_t, \pi_t) = -cu_t + \min_{S_t \geq u_t} \{cS_t + G_t(S_t|\pi_t, l) + E_{w_t|\pi_t}[J_{t+1}(S_t - w_t, T(\pi_t|w_t))]\},$$

$$t = 0, \dots, T. \quad (3.8)$$

It is important to characterize the cost functions, $J_t(u_t, \pi_t)$ and the optimal policy. We first prove the convexity of the cost function.

Theorem 1. $J_t(u_t, \pi_t)$ is a convex function of u_t for all π_t .

Proof: The proof will proceed inductively. First, note that $J_{T+1}(u_{T+1}, \pi_{T+1}) = -cu_{T+1}$ is a convex function of u_{T+1} . Next, assume that $J_{t+1}(u_{t+1}, \pi_{t+1})$ is convex in u_{t+1} , then

we will show that $J_t(u_t, \pi_t)$ is also convex in u_t . Note that $G_t(y_t|\pi_t, l)$ is convex in y_t and that $\lim_{|y_t| \rightarrow \infty} G_t(y_t|\pi_t, l) = \infty$. Also, note that because $J_{t+1}(u_{t+1}, \pi_{t+1})$ is convex in u_{t+1} , $E_{w_t|\pi_t}[J_{t+1}(y_t - w_t, T(\pi_t|w_t))]$ is also convex in y_t . Therefore,

$$H_t(y_t|\pi_t, l) = cy_t + G_t(y_t|\pi_t, l) + E_{w_t|\pi_t}[J_{t+1}(y_t - w_t, T(\pi_t|w_t))]$$

is convex in y_t . Define $S_t(\pi_t)$ as the value of y_t that minimizes $H_t(y_t|\pi_t, l)$. If $u_t < S_t(\pi_t)$, then we have that

$$J_t(u_t, \pi_t) = c(S_t(\pi_t) - u_t) + G_t(S_t(\pi_t)|\pi_t, l) + E_{w_t|\pi_t}[J_{t+1}(S_t(\pi_t) - w_t, T(\pi_t|w_t))]$$

which is a non-increasing linear function of u_t . However, if $u_t \geq S_t(\pi_t)$ then

$$J_t(u_t, \pi_t) = G_t(u_t|\pi_t, l) + E_{w_t|\pi_t}[J_{t+1}(u_t - w_t, T(\pi_t|w_t))]$$

which is non-decreasing and convex in u_t . Because $J_t(u_t, \pi_t)$ is non-increasing and convex for $u_t < S_t(\pi_t)$ and non-decreasing and convex for $u_t \geq S_t(\pi_t)$, it is convex for all u_t . \square

This result demonstrates that a state-dependent base stock policy is optimal for this problem. That is, there exists a value $S_t(\pi_t)$ that minimizes the function $H_t(y_t|\pi_t, l)$ such that the optimal policy is

$$a_t = \begin{cases} S_t(\pi_t) - u_t, & \text{if } u_t < S_t(\pi_t) \\ 0, & \text{if } u_t \geq S_t(\pi_t) \end{cases}$$

This result allows us to restrict the search for an optimal control policy to the class of base stock policies and greatly reduces the computational requirements. This result also justifies the use of base stock control in suboptimal control policies. Although the optimal policy has a relatively simple structure, it is computationally difficult to compute the policy because the stock level is state-dependent. Therefore, we describe several possible suboptimal control policies in the next section that can be computed much more quickly.

3.4 Suboptimal Control

In this section we discuss suboptimal control policies for this problem. In describing these models, we assume a positive lead time and define the cost-to-go function using inventory position as in Equation (3.8). Because of the intense computational requirements for computing optimal policies, it is valuable to develop and study suboptimal control policies that can be computationally feasible for practical application. There are many different forms of suboptimal control and their effectiveness depends a great deal on the problem structure and parameters. This section describes several different suboptimal control strategies for the basic model described in Section 3.3. We assume that a control policy will always use feedback on the current inventory state variable, x_t or u_t , to determine the order quantity, a_t . However, the policy may or may not use feedback to revise the prior distribution, π_t . Therefore, in the following discussion, the terms *open loop*, *closed loop*, and *feedback* are used to indicate whether or not a control policy uses an updated prior, π_t , to select the control action in period t .

3.4.1 Certainty Equivalence Control

In this approach, some function of the information vector, I_t , is used to generate an estimate, \hat{d}_t , of d_t . Assuming that $d_t = \hat{d}_t$, a complete information problem is solved to determine the initial base stock level S_0 . In practice, this process is repeated at the beginning of each period using the newly updated information vector. This approach is essentially the one used by most inventory control systems. The estimate \hat{d}_t is computed using some type of forecasting model, then depending on the problem either a stationary or non-stationary demand model is used to compute S_0 . This approach ignores the uncertainty in the estimate \hat{d}_t and does not account for the use of feedback about d_{t+n} in future periods when calculating the current solution.

In each of the three CEC policies, we use a maximum likelihood estimate (MLE) to generate \hat{d}_t . For the first policy (CEC 1), the estimates are:

$$\hat{d}_t = \operatorname{argmax}_i \{\pi_{it}\} \quad (3.9)$$

$$\hat{d}_{t+n} = \operatorname{argmax}_j \{P_{\hat{d}_{t+n-1}, j}\}, \quad n = 1, \dots, T - t. \quad (3.10)$$

The estimate at time t is the distribution that is most likely based on the current prior. In subsequent periods, $s > t$, the estimate, \hat{d}_s , is the distribution that is the most likely given \hat{d}_{s-1} .

For the second CEC policy (CEC 2), the estimates are:

$$\hat{d}_t = \operatorname{argmax}_i \{\pi_{it}\} \quad (3.11)$$

$$\hat{d}_{t+n} = \operatorname{argmax}_j \left\{ \sum_{i=1}^N \pi_{ij}(P_{ij}^n) \right\}, \quad n = 1, \dots, T-t. \quad (3.12)$$

The initial estimate is the same as in CEC 1. However, subsequent estimates are a function of both the initial prior and the n-step transition matrix. The cost-to-go function used to compute the stock level for a period, t , given the estimators $\hat{d}_t, \hat{d}_{t+1}, \dots, \hat{d}_T$, and the current inventory position, u_t , is

$$\begin{aligned} \bar{J}_{t+n}^C(u_{t+n}) = & -cu_{t+n} + \min_{S_{t+n} \geq u_{t+n}} \left\{ cS_{t+n} + G_{t+n}(S_{t+n} | \hat{d}_{t+n}, l) \right. \\ & \left. + E_{w_{t+n} | \hat{d}_{t+n}} [\bar{J}_{t+n+1}^C(S_{t+n} - w_{t+n})] \right\}, \quad n = 0, 1, \dots, T-t. \end{aligned} \quad (3.13)$$

where G_{t+n} is calculated from Equation(3.6) using \hat{d}_{t+n} instead of π_{t+n} . Note that Equation (3.13) is the recursion for the nonstationary, complete information problem where d_t is known.

In the above two CEC policies, the estimate at any given time is one of the N discrete distributions. In the third CEC policy (CEC 3), the estimate at time t , \hat{d}_t , is the same as above, but subsequent distributions are conditional distributions based on \hat{d}_t and the n-step transition matrix, P^n . The distribution in subsequent periods is

determined by the following:

$$P[w_{t+n} = k | \hat{d}_t] = \sum_{j=1}^N P_{\hat{d}_t, j}^n r_{jk}.$$

The cost-to-go function for CEC 3 is

$$\begin{aligned} \bar{J}_{t+n}^{C3}(u_{t+n}) = & -cu_{t+n} + \min_{S_{t+n} \geq u_{t+n}} \left\{ cS_{t+n} + G_{t+n}(S_{t+n} | \hat{d}_t, l) \right. \\ & \left. + E_{w_{t+n} | \hat{d}_t} [\bar{J}_{t+n+1}^{C3}(S_{t+n} - w_{t+n})] \right\}, \quad n = 0, 1, \dots, T-t. \end{aligned} \quad (3.14)$$

This procedure is similar to the Open-Loop Feedback Control discussed next, except that the demand distributions are conditional on the maximum likelihood estimate, \hat{d}_t , rather than π_t .

3.4.2 Open-Loop Feedback Control

Open-loop strategies do not use feedback to update estimates of π_t for $t > 0$. A decision is made at time zero for an initial order-up-to level under the assumption that no feedback will be used in the future to calculate new stock levels. However, in practice, an open-loop control problem may be solved each period with the previous demand observation used as feedback to update the current prior. This type of control may be referred to as Open-Loop Feedback Control (OLFC) (see Bertsekas [5]). The transition equation $\pi_t = T(\pi_{t-1} | w_{t-1})$ updates the prior at the beginning of t , and the cost-to-go function

$$\begin{aligned} \bar{J}_{t+n}^O(u_{t+n}) = & -cu_{t+n} + \min_{S_{t+n} \geq u_{t+n}} \left\{ cS_{t+n} + G_{t+n}(S_{t+n}|\pi_t, l) \right. \\ & \left. + E_{w_{t+n}|\pi_t} [\bar{J}_{t+n+1}^O(S_{t+n} - w_{t+n})] \right\}, \quad n = 0, 1, \dots, T-t, \end{aligned} \quad (3.15)$$

provides the new base stock level, $S_t(\pi_t)$. This approach takes advantage of both open and closed-loop control. The equation for \bar{J}^O assumes that feedback will not be used in future periods (open loop) and thereby greatly reduces the computational requirements. However, because π_t is updated using the prior observations (closed loop), it does incorporate the information available at t to characterize the distribution of w_t . Another way of describing this approach is that Equation (3.15) is implemented on a “rolling-horizon” basis.

3.4.3 Limited Look-Ahead Control

This class of suboptimal control policies optimizes the dynamic problem for only a limited amount of time into the future, say L periods. The appropriate transition equation, $\pi_t = T(\pi_{t-1}|w_{t-1})$ updates the prior at the beginning of t . When $L = 0$, we refer to the policy as a *myopic* policy and the approximate cost equation is defined as

$$\bar{J}_t^{L0}(u_t, \pi_t) = \min_{S_t \geq u_t} \{c(S_t - u_t) + G_t(S_t|\pi_t, l)\}, \quad (3.16)$$

that is, the base stock level is chosen to minimize the costs only for the current period, t .

For $L = 1$, we refer to the policy as the single period look-ahead policy and the equation is

$$\bar{J}_t^{L1}(u_t, \pi_t) = \min_{S_t \geq u_t} \{c(S_t - u_t) + G_t(S_t | \pi_t, l) + E_{w_t | \pi_t} [\bar{J}_{t+1}^{L0}(S_t - w_t, T(\pi_t | w_t))]\}. \quad (3.17)$$

For $L = 2$, we refer to the policy as the two period look-ahead policy and the equation is

$$\bar{J}_t^{L2}(u_t, \pi_t) = \min_{S_t \geq u_t} \{c(S_t - u_t) + G_t(S_t | \pi_t, l) + E_{w_t | \pi_t} [\bar{J}_{t+1}^{L1}(S_t - w_t, T(\pi_t | w_t))]\}. \quad (3.18)$$

It is possible to look even further ahead into the future; however, the computational requirements increase exponentially as the look-ahead period increases. Note that to solve a 1-period look-ahead problem, the second period solution is a myopic policy. Similarly, in a 2-period look-ahead policy, the second period solution is a single period policy. The efficacy of these strategies and the best choice of L will depend upon the dynamics of the core process as determined by the transition matrix P .

Dist.	Initial		Truncated
	Mean	Std. Dev.	Mean
1	1.00	1.01	1.00
2	9.00	3.01	8.97
3	16.00	4.01	14.20

Table 3.2: Demand Distributions

3.5 Experiment Design

In this section, we compare the results of an optimal policy to the seven suboptimal policies described in the previous section. There are three discrete distributions that form the set of possible distributions. Each of the distributions were initially generated using negative binomial distributions and then truncated at a maximum 18 and rescaled. Table 3.2 shows the demand parameters. We list the demand distributions in order of increasing mean demand. The second and third columns show the mean and standard deviation before truncation, and the fourth column shows the mean of the truncated distributions. Figure 3.2 shows the three truncated distributions. In Figure 3.3 we show the conditional probability of each distribution after a single observation assuming a uniform prior at time 0. If the observation is between 0 – 3, 5 – 9, or 15 – 18 then there is a high probability (above 80%) that the current state is distribution 1, 2, or 3 respectively. Similarly, a demand of 11, 12 or 13 will not provide as much information about whether the distribution is 2 or 3. Therefore, a single observation may or may not provide a strong indication of the current state.

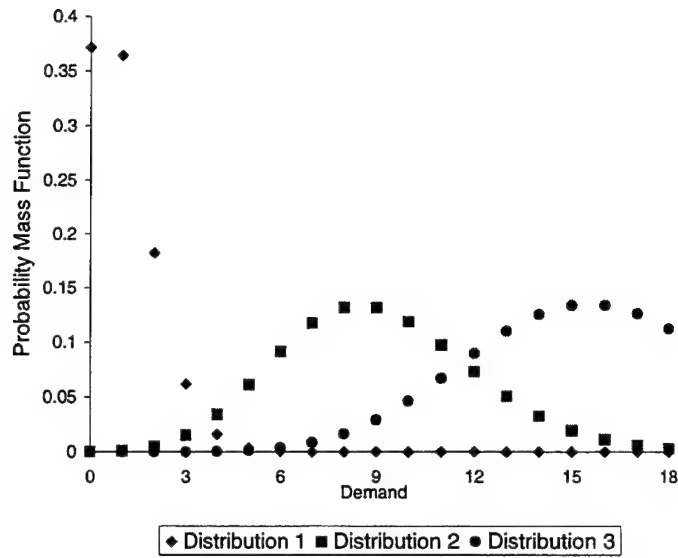


Figure 3.2: Discrete Demand Distributions

In our example we model backordering and unit cost = 0. Table 3.3 lists the parameters for our experimental design. For each of the three lead times (0,1, and 2), our experimental design included 84 different runs. For all of the following results, the planning horizon is 5 periods. Any longer horizon length makes the computational requirements for computing an optimal policy excessively long, especially with a lead time of 2. Procurement cost is set at 0, and the lead time ranges from 0 to 2. We examine 4 different priors for the distribution in effect at time 0. In the first case, the distribution is most likely in the high distribution, 3. In the second case, the distribution is most likely the low, 1. The third prior indicates that there is no strong information about the current demand distribution. Finally, the last prior indicates a strong belief that the

Periods	=	5		
Leadtime	=	0	1	2
Critical Ratio	=	0.5	0.65	0.8
Prior 1 [$\delta(\pi(t) = .4)$]	=	0.1	0.3	0.6
Prior 2 [$\delta(\pi(t) = .4)$]	=	0.6	0.3	0.1
Prior 3 [$\delta(\pi(t) = .0)$]	=	0.33	0.34	0.33
Prior 4 [$\delta(\pi(t) = .7)$]	=	0.1	0.8	0.1
Tran 1		0.8	0.1	0.1
[erg(P)=.7]	=	0.1	0.8	0.1
		0.1	0.1	0.8
Tran 2		0.2	0.4	0.4
[erg(P)=.2]	=	0.4	0.2	0.4
		0.4	0.4	0.2
Tran 3		0.333	0.333	0.333
[erg(P)= 0]	=	0.333	0.333	0.333
		0.333	0.333	0.333
Tran 4		0.8	0.1	0.1
[erg(P)=.9]	=	0.0	0.8	0.2
		0.0	0.0	1.0
Tran 5		0.2	0.4	0.4
[erg(P)=.6]	=	0.0	0.2	0.8
		0.0	0.0	1.0
Tran 6		1.0	0.0	0.0
[erg(P)=.9]	=	0.2	0.8	0.0
		0.1	0.1	0.8
Tran 7		1.0	0.0	0.0
[erg(P)=.6]	=	0.8	0.2	0.0
		0.4	0.4	0.2

Table 3.3: Experiment Parameters

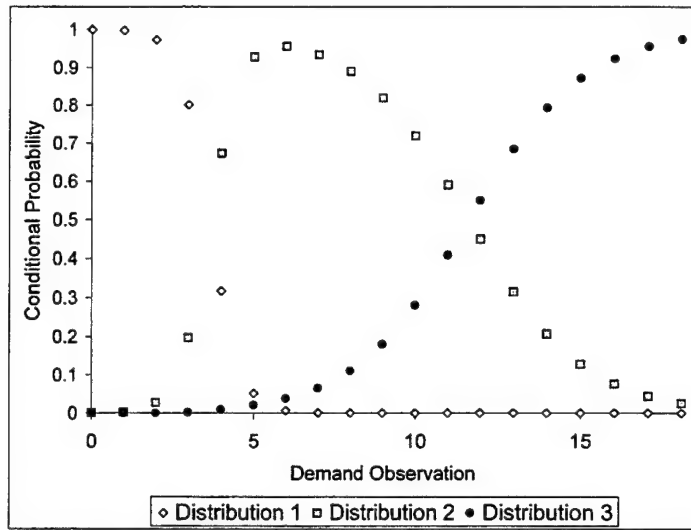


Figure 3.3: Conditional State Probability after Demand Observation

distribution is the middle one, 2. We also show for each prior the values for

$$\delta(\pi_t) = \frac{N}{2(N-1)} \sum_{i=1}^N \left| \pi_{it} - \frac{1}{N} \right|$$

which ranges from 0 to 1. The function δ measures the dispersion of the prior distribution π_t from a discrete uniform distribution on $\{1, \dots, N\}$. As δ approaches 1, there is a high confidence that a specific distribution is in effect. Similarly, as δ approaches 0, there is a low confidence that a specific distribution is in effect. We experiment with 7 transition matrices. The first one represents a very stable environment and the second represents a highly unstable environment. For the third transition matrix, the transitions all have equal probability. Matrices four and five represent a movement to a permanently high level of demand, and the last 2 matrices represent movements to a permanently low

level of demand. We examine three critical values: .50, .65, and .80. These ratios were determined by keeping the holding cost constant and adjusting the penalty cost such that $CR = \frac{p}{p+h}$. A value close to one indicates a high relative penalty cost or, equivalently, a high service level. It should be noted that in our problem, there is no penalty for underestimating the demand in subsequent periods. The reason for this is that there are no capacity constraints, so if the current inventory level (position) is below the order-up-to level, the order-up-to level can always be achieved. On the other hand, it is possible to move into a subsequent period with too much inventory. It is in these instances that a suboptimal policy is penalized. It would not be difficult to modify this problem to penalize a situation when a large amount of inventory must be produced at once. This could be done by adding capacity constraints or convex purchase costs. Finally, under the CEC policies, when there is a tie between 2 distributions most likely to occur, we choose distribution 2.

To compare the policies, we adopted the following procedure. We first compute the optimal policy for each parameter set. We then compare the optimal expected cost to the actual expected cost of each suboptimal policy. Let σ be a suboptimal policy with state dependent stock levels, $S_t^\sigma(\pi_t)$, and order quantities, $a_t^\sigma(u_t, \pi_t) = \max\{0, S_t^\sigma(\pi_t) - u_t\}$. Then the actual expected cost of σ can be computed using the recursion

$$J_t^\sigma(u_t, \pi_t) = c \cdot \max\{0, S_t^\sigma(\pi_t) - u_t\} + G_t(\max\{S_t^\sigma(\pi_t), u_t\} | \pi_t, l) \\ + E_{w_t | \pi_t} [J_{t+1}^\sigma(\max\{S_t^\sigma(\pi_t), u_t\} - w_t, T(\pi_t | w_t))] , \quad t = 0, \dots, T. \quad (3.19)$$

Thus, $J_0^\sigma(u_0, \pi_0)$ is the expected cost of policy σ over the given horizon. The calculation requirements to compute the expected costs are almost as intensive as they are to compute the optimal solution. However, it is unnecessary to compute the expected costs in practice.

Lead time	Optimal	Myopic	LA - 1	LA - 2	OLFC	CEC 2
0	1m 24s	<1s	<1s	1s	<1s	<1s
1	7m 54s	<1s	<1s	5s	<1s	<1s
2	1h 47m 07s	<1s	5s	1m 25s	2s	2s

Table 3.4: User Runtime

The run times on a SUN workstation to compute six policies for one set of representative parameters are shown in Table 3.4. The optimal column shows the time to solve one 5-period problem optimally. The next five columns show the time to solve the problem once at time period zero only, multiplied by five to account for the five decisions that are made over the entire horizon. These columns show times for the myopic, 1-period look-ahead, 2-period look-ahead, open-loop feedback control, and the second certainty equivalent control policies. The calculation of the exact expected costs required substantially more time, *e.g.*, 2.10 hours of user time on the same system for a 2-period look-ahead problem with a lead time of 2.

3.6 Results and Analysis

The complete results are presented in Tables 3.11 through 3.19 of the appendix. For emphasis, values of "0" are indicated by blank cells. The first three columns designate

the parameters of the problem instance. The next column shows the optimal expected cost for the 5-period problem. The remaining columns show the percentage by which each suboptimal policy exceeds the optimal cost. Other tables are organized likewise. The most significant result is that the full consideration of the uncertainty of the demand distribution, even over a limited horizon, leads to substantial improvements over the CEC policies. Even the myopic policy performs substantially better than the CEC policies.

Table 3.5 shows the percentage of cases in which each suboptimal policy achieves the best cost among these suboptimal policies. In some cases, two or more policies tied for the best suboptimal cost. Table 3.6 shows the average percentage that each suboptimal

2 LA	1 LA	OLFC	Myopic	CEC 3	CEC 2	CEC 1
97.62%	50.00%	44.84%	43.25%	1.59%	1.59%	1.19%

Table 3.5: Percent of Cases Achieving Best Suboptimal Cost

policy exceeded the optimal cost. In this and other summary tables, the percentages reported are the average deviation of each particular case from its respective optimal cost. The results in Tables 3.5 and 3.6 lead us to observe that the best suboptimal controls based on cost are, in order from best to worst, 2 LA, 1 LA, OLFC, Myopic, CEC 3, CEC 2, and CEC 1. OLFC and the myopic controls are fairly close in their performance. The look-ahead, myopic, and OLFC control policies perform much better than the CEC policies.

We summarize the results for the three lead times in Table 3.7. We first see that the optimal cost goes up as the lead time increases. This is certainly expected because

2 LA	1 LA	OLFC	Myopic	CEC 3	CEC 2	CEC 1
0.33%	1.53%	2.03%	5.46%	12.56%	16.68%	18.04%

Table 3.6: Average Percent Above Optimal Cost

Lead time	Cost	Myopic	1LA	2LA	OLFC	CEC1	CEC2	CEC3
0	\$252	5.03	1.31	0.25	1.79	14.68	14.08	14.58
1	\$370	5.74	1.61	0.37	2.12	17.16	14.32	11.77
2	\$456	5.60	1.66	0.38	2.17	22.27	21.64	11.33

Table 3.7: Summary of Results by Lead Time

demand uncertainty increases as the lead time increases. Note that the cost appears to be a concave function of the lead time. The three look-ahead policies (Myopic, 1 LA, 2 LA) perform quite well in many instances. Even the myopic policy averages between 5 and 6 percentage points above the optimal. For the 1-period look-ahead policy, the average percentage above optimal is about one quarter of that of the myopic policy. The consideration of only one additional period provides substantial improvement over a fairly good myopic policy. The 2-period look-ahead policies are less than one quarter of the percentages for the 1-period look-ahead. The 2-period look-ahead policies are typically at least 30-50 times less costly than the best CEC policy. Examination of the detailed results in the appendix shows that, as expected, the 1-period look-ahead is always at least as good as the myopic policy and the 2-period look-ahead policy is always as good as the other two look-ahead policies.

The OLFC average performance is comparable to the 1-period look-ahead policy. The OLFC never performs worse than the myopic policy, although it is usually only

marginally better. This is because the OLFC does not anticipate feedback to revise the prior distribution in a given solution, but it does consider, unlike the myopic policy, the uncertainty in the future due to transitions. Although the OLFC outperformed the myopic policy, it never outperformed the 1-period look-ahead policy. This result suggests the advantage of anticipating feedback to solve the problem. However, given that the OLFC may be much faster than the look-ahead policies, it may still be very useful in practice. Also, a composite look-ahead and OLFC policy might provide excellent performance. For example, a 1-period look-ahead policy might be used with an OLFC for the remainder of the horizon.

Most importantly, the CEC policies rarely outperform the look-ahead or OLFC policies. All of the average CEC percentages were at least 11% above the optimal cost. CEC 3 generally outperforms the other two CEC policies, sometimes very significantly. The improvement in CEC 3 relative to CEC 1 and CEC 2 increases as the lead time increases. This is because after time t the future demand distribution is a composite distribution based on the original prior as well as the transition matrix. Because these or similar policies are often used in practice, these results suggest that even a very simple policy, such as the myopic policy, that fully considers the demand uncertainty will perform significantly better than a CEC policy.

We used three different critical ratios (see Table 3.8). We see that the optimal costs go up as the critical ratio increases. This is as expected, because we modified the critical ratio by keeping the holding cost constant and increasing the penalty cost. This

C. R.	Cost	Myopic	1LA	2LA	OLFC	CEC1	CEC2	CEC3
0.50	\$263	5.79	1.31	0.24	2.54	17.12	14.02	12.46
0.65	\$351	5.76	1.51	0.36	2.02	14.96	13.18	10.78
0.80	\$463	4.81	1.76	0.40	1.52	22.04	22.83	14.44

Table 3.8: Summary of Results by Critical Ratio

equates to the practice of requiring a higher service level as the penalty for non-service increases. The optimal cost appears to be a convex function of the critical ratio. Again, the 2-period look-ahead policy performs extremely well. The CEC policies generally perform poorly. CEC 3 outperforms CEC 1 and CEC 2 much more at higher critical ratios. However, if one keeps all parameters constant and changes only the critical ratio, the performance of any given suboptimal policy relative to another cannot be predicted. For example, in Table 3.11 the CEC policies improve with prior 1 and transition matrix 1 as the critical ratio increases, but worsen as the critical ratio increases with prior 2 and transition matrix 1.

The results for the seven transition matrices are summarized in Table 3.9. Each of

P	erg(P)	Cost	Myopic	1LA	2LA	OLFC	CEC1	CEC2	CEC3
1	0.7	\$411	1.41	.35	.07	1.36	20.19	20.19	11.17
2	0.2	\$445					30.85	36.82	24.49
3	0.0	\$447					4.49	4.49	3.08
4	0.8	\$320	1.48	0.39	0.09	1.48	24.76	24.38	13.84
5	0.7	\$242					6.61	4.53	5.43
6	0.9	\$410	12.95	4.12	1.05	10.00	25.36	18.28	17.28
7	0.6	\$237	22.36	5.82	1.12	1.35	14.00	8.06	12.61

Table 3.9: Summary of Results by Transition Matrix

the non-CEC policies achieve the optimal cost (or very close to it) for transition matrices 2, 3, and 5. These matrices represent situations when the transitions are highly random ($P = 2, 3$) or rapidly moving to a permanently high demand state ($P = 5$). The myopic policy does very poorly when the transitions are to a lower demand level ($P = 6, 7$) because there is a very strong likelihood of carrying excess inventory forward. The CEC policies perform very poorly when the transition process is highly unstable ($P = 2$). This observation leads us to consider the following measure of *ergodicity* for a transition matrix P (see Hopp, Bean, and Smith [14])

$$\text{erg}(P) = \max_{i,j} \sum_s \frac{1}{2} |p_{is} - p_{js}|.$$

This quantity is a measure of how quickly the Markov process converges to its stationary distribution. When $\text{erg}(P)$ is close to 1, convergence is slow, and when it is close to zero the rate of convergence is fast. Table 3.9 also reports the ergodicity values for each transition matrix. We see that the open-loop feedback control policies appear to perform better as the ergodicity measure decreases ($P = 2, 3$). It appears also to do very well for $P = 5$, a fast transition to a high demand, but this is due to the fact that there is no penalty for going into a future period under-stocked. Lower ergodicity values usually, but not necessarily, indicate better performance for limited look-ahead policies also. When the ergodicity value is high, knowledge of the current state of the core process is much more valuable in predicting subsequent demands. Therefore, adaptive control policies

that update π_t using the current observation have greater benefit over non-adaptive policies for problems with a higher value of $\text{erg}(P)$. For example, with transition matrix 6 ($\text{erg}(P) = .9$), the myopic policy is 12.95% but improves drastically to 4.12% and 1.05% with the other two look-ahead policies. Furthermore, suboptimal policies will not perform as well relative to optimal policies for higher values of $\text{erg}(P)$ and such problems will provide more distinction in performance among suboptimal policies.

Pri.	$\delta(\pi(t))$	Cost	Myo.	1LA	2LA	OLFC	CEC1	CEC2	CEC3
1	0.4	\$371	3.55	1.18	0.36	0.87	19.36	15.99	14.85
2	0.4	\$344	7.61	1.42	0.21	3.03	25.59	24.30	15.28
3	0.0	\$380	8.06	2.60	0.58	3.43	14.88	15.14	12.02
4	0.7	\$341	2.60	0.91	0.18	0.78	12.31	11.28	8.08

Table 3.10: Summary of Results by the Prior

Table 3.10 summarizes the results for the four different priors. The suboptimal policies generally perform the worst when the $\delta(\pi_t)$ is low, because there is little confidence in the belief of which distribution is in effect when $\delta(\pi_t)$ is low. The poorest results for the look-ahead policies are usually obtained when $\delta(\pi_t) = 0$. There are some exceptions, for example, in Table 3.13 the myopic policy does very poorly when there is a transition to a low demand distribution ($P = 6, 7$) even if there is a strong belief that the demand is low (prior 2). This is because if the demand is actually in a higher range, there is a significant penalty to pay for under-stocking. The 1-period look-ahead policy accounts for this and substantially improves the results as it does in other areas. The worst case for the policies that explicitly consider the prior distribution is prior 3 which indicates

no current information is available. Naturally, when information is available these policies will improve. However, the worst case prior for the CEC policies is prior 2, which indicates a strong belief that the demand is low. This is because the CEC policies ignore the possibility that the demand distribution may not be low at time 0 and, therefore, they under-stock and incur high penalty costs.

3.7 Conclusion

We have demonstrated that a policy that uses feedback and fully accounts for the demand uncertainty in this problem can substantially outperform a certainty equivalent policy. Often times in practice, managers feel compelled to adopt CEC policies, although this is frequently not necessary and often very costly. By incorporating more of the actual uncertainty of the demand process even over a limited horizon into a suboptimal policy can provide dramatic improvements in the solution of this problem. In many cases even a myopic policy will provide significantly better results. Additionally, Open-Loop Feedback Control policies also will often perform exceptionally well when implemented in a "rolling horizon" fashion. Look-ahead policies should also be tractable for many realistic problems.

The most effective policies consider the uncertainty of the demand distribution through an *a priori* distribution. This prior can be estimated through many different means and some future research can focus on efficient ways to obtain these priors as well as evaluate the sensitivity of the solution to the accuracy of the prior.

3.A Results

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	1	0.50	\$217	1.10	0.62	0.03	0.99	13.14	13.14	13.14
1	1	0.65	\$269	0.53	0.36	0.27	0.52	7.81	7.81	7.81
1	1	0.80	\$324	0.17	0.14	0.14	0.17	6.37	6.37	6.37
2	1	0.50	\$178	0.43	0.11	0.01	0.39	3.83	3.83	3.83
2	1	0.65	\$263	3.37	0.30	0.01	3.37	6.19	6.19	6.19
2	1	0.80	\$384	0.58	0.01		0.58	18.23	18.23	18.23
3	1	0.50	\$216	7.98	1.32	0.29	7.92	12.29	12.29	12.29
3	1	0.65	\$291	2.31	0.79	0.02	2.31	3.54	3.54	3.54
3	1	0.80	\$371	0.81	0.02		0.81	4.93	4.93	4.93
4	1	0.50	\$190	1.21	0.87	0.02	1.19	3.74	3.74	3.74
4	1	0.65	\$248	0.21	0.03		0.19	2.03	2.03	2.03
4	1	0.80	\$321	0.07			0.06	4.96	4.96	4.96
1	2	0.50	\$245					21.05	18.33	21.05
1	2	0.65	\$312					40.60	40.17	40.60
1	2	0.80	\$381					86.10	88.83	86.10
2	2	0.50	\$244					20.93	18.56	20.93
2	2	0.65	\$323					39.35	38.58	39.35
2	2	0.80	\$404					86.25	87.48	86.25
3	2	0.50	\$257					19.35	17.50	19.35
3	2	0.65	\$328					39.19	38.56	39.19
3	2	0.80	\$402					85.72	86.72	85.72
4	2	0.50	\$241					20.93	19.37	20.93
4	2	0.65	\$308					44.56	44.04	44.56
4	2	0.80	\$381					96.45	97.28	96.45
1	3	0.50	\$249					3.19	3.19	3.19
1	3	0.65	\$314					3.36	3.36	3.36
1	3	0.80	\$382					13.97	13.97	13.97
2	3	0.50	\$249					0.47	0.47	0.47
2	3	0.65	\$329					5.17	5.17	5.17
2	3	0.80	\$410					26.46	26.46	26.46
3	3	0.50	\$261							
3	3	0.65	\$332					1.97	1.97	1.97
3	3	0.80	\$405					16.18	16.18	16.18
4	3	0.50	\$241							
4	3	0.65	\$308					1.70	1.70	1.70
4	3	0.80	\$381					14.29	14.29	14.29

Table 3.11: Results: Lead time=0; Trans Matrices 1, 2 & 3

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	4	0.50	\$157	0.61	0.61		0.61	11.44	11.44	11.44
1	4	0.65	\$198	0.30	0.30	0.30	0.30	7.60	7.60	7.60
1	4	0.80	\$241	0.16	0.16	0.16	0.16	7.75	7.75	7.75
2	4	0.50	\$154	0.15	0.01		0.15	3.06	3.06	3.06
2	4	0.65	\$234	3.64	0.29	0.01	3.64	7.43	7.43	7.43
2	4	0.80	\$348	0.63			0.63	22.18	22.18	22.18
3	4	0.50	\$175	9.26	1.44	0.33	9.26	11.73	11.73	11.73
3	4	0.65	\$242	2.49	0.84	0.01	2.49	3.47	3.47	3.47
3	4	0.80	\$312	0.89			0.89	6.97	6.97	6.97
4	4	0.50	\$156	0.73	0.69		0.73	4.39	4.39	4.39
4	4	0.65	\$206	0.02			0.02	5.34	5.34	5.34
4	4	0.80	\$268					13.31	13.31	13.31
1	5	0.50	\$135					6.35	6.35	6.35
1	5	0.65	\$167					3.68	3.68	3.68
1	5	0.80	\$200					1.78	1.78	1.78
2	5	0.50	\$152					9.03	9.03	9.03
2	5	0.65	\$203					10.30	10.30	10.30
2	5	0.80	\$254					24.14	24.14	24.14
3	5	0.50	\$155					4.68	4.68	4.68
3	5	0.65	\$195					3.35	3.35	3.35
3	5	0.80	\$236					6.82	6.82	6.82
4	5	0.50	\$131					2.75	2.75	2.75
4	5	0.65	\$166					1.57	1.57	1.57
4	5	0.80	\$205					1.50	1.50	1.50

Table 3.12: Results: Lead time=0; Trans Matrices 4 & 5

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	6	0.50	\$229	4.99	3.09	0.94	4.66	20.65	20.65	20.65
1	6	0.65	\$287	2.54	1.64	0.27	2.50	11.01	11.01	11.01
1	6	0.80	\$348	1.35	1.04	0.01	1.31	5.85	5.85	5.85
2	6	0.50	\$123	3.80	1.61	0.21	3.57	8.54	8.53	8.54
2	6	0.65	\$179	21.98	1.31	0.10	6.01	6.47	6.47	6.47
2	6	0.80	\$272	29.07	7.81	0.60	21.01	6.61	6.61	6.61
3	6	0.50	\$192	27.81	4.87	0.24	20.92	34.78	28.10	34.78
3	6	0.65	\$270	16.75	7.72	2.01	16.65	15.47	11.70	15.47
3	6	0.80	\$371	8.51	5.12	2.45	8.49	2.91	2.02	2.91
4	6	0.50	\$186	7.57	2.27	0.75	7.45	14.83	14.64	14.83
4	6	0.65	\$246	3.01	1.95	0.82	2.61	6.52	6.39	6.52
4	6	0.80	\$319	2.20	1.38	0.69	2.18	1.94	1.85	1.94
1	7	0.50	\$147	19.32	4.69	1.50	1.69	35.22	16.50	24.86
1	7	0.65	\$218	19.54	4.99	0.67	1.10	21.69	10.43	21.69
1	7	0.80	\$322	15.28	4.28	1.30	1.58	16.58	13.47	16.58
2	7	0.50	\$84	2.11	0.57		0.57	0.16	0.16	0.16
2	7	0.65	\$136	30.70	1.16		0.95	5.05	5.05	5.05
2	7	0.80	\$232	40.69	9.73	0.81	0.94	8.46	8.46	8.46
3	7	0.50	\$128	44.35	1.81		1.82	28.99	28.99	28.99
3	7	0.65	\$206	31.14	11.72	2.80	2.99	13.26	13.26	13.26
3	7	0.80	\$328	21.27	10.09	2.11	2.39	6.43	6.43	7.85
4	7	0.50	\$100	9.57	2.66			2.74	2.74	2.74
4	7	0.65	\$148	6.78	4.53	0.58	0.65	1.84	1.84	1.84
4	7	0.80	\$226	14.34	4.85	0.70	0.83	2.01	2.01	2.60

Table 3.13: Results: Lead time=0; Trans Matrices 6 & 7

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	1	0.50	\$341	1.69	0.47	0.25	1.58	28.80	28.80	21.69
1	1	0.65	\$429	0.72	0.49	0.01	0.67	15.50	15.50	10.25
1	1	0.80	\$529	0.33	0.19	0.12	0.32	11.62	11.62	6.86
2	1	0.50	\$310	1.18	0.19	0.12	1.15	15.45	15.45	13.91
2	1	0.65	\$414	0.33	0.18		0.33	34.09	34.09	26.32
2	1	0.80	\$541	0.24	0.04		0.24	75.68	75.68	45.24
3	1	0.50	\$335	3.19	0.51	0.02	3.12	17.17	17.17	14.40
3	1	0.65	\$435	2.01	0.67	0.04	2.00	7.41	7.41	6.82
3	1	0.80	\$560	2.55	0.46		2.55	5.25	5.25	3.21
4	1	0.50	\$297	0.79	0.31	0.22	0.73	7.37	7.37	5.93
4	1	0.65	\$383	0.73	0.10	0.01	0.72	3.79	3.79	2.94
4	1	0.80	\$490	0.63	0.04	0.01	0.62	5.30	5.30	2.46
1	2	0.50	\$335					16.87	8.31	8.75
1	2	0.65	\$447				0.01	21.65	16.18	13.20
1	2	0.80	\$583					35.89	43.04	18.81
2	2	0.50	\$335					17.09	7.39	8.92
2	2	0.65	\$447				0.02	24.26	13.44	13.78
2	2	0.80	\$587					40.23	33.69	18.89
3	2	0.50	\$340					15.64	11.58	8.41
3	2	0.65	\$455				0.01	21.18	22.99	12.85
3	2	0.80	\$594					36.73	51.16	17.86
4	2	0.50	\$336					15.89	8.64	8.98
4	2	0.65	\$448				0.01	23.42	16.26	13.70
4	2	0.80	\$582					41.36	38.05	19.91
1	3	0.50	\$338					4.36	4.36	1.79
1	3	0.65	\$450					0.74	0.74	1.82
1	3	0.80	\$584	0.01				4.51	4.51	1.57
2	3	0.50	\$338					3.21	3.21	1.54
2	3	0.65	\$451					3.52	3.52	2.47
2	3	0.80	\$592					11.76	11.76	3.63
3	3	0.50	\$342					2.70	2.70	0.63
3	3	0.65	\$458							0.33
3	3	0.80	\$597					5.75	5.75	0.63
4	3	0.50	\$334					2.39	2.39	0.51
4	3	0.65	\$445					0.15	0.15	0.27
4	3	0.80	\$578					5.83	5.83	0.52

Table 3.14: Results: Lead time=1; Trans Matrices 1, 2 & 3

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	4	0.50	\$227	1.26	0.38	0.38	1.26	18.01	17.20	17.31
1	4	0.65	\$288	1.11	0.25	0.25	1.11	12.21	10.66	10.17
1	4	0.80	\$361	0.14	0.14	0.14	0.14	11.88	9.55	8.88
2	4	0.50	\$267	1.35	0.31		1.35	17.66	16.68	16.34
2	4	0.65	\$360	0.39			0.39	46.63	44.96	35.05
2	4	0.80	\$474	0.53			0.53	100.24	97.84	60.49
3	4	0.50	\$260	3.43	0.52		3.43	11.03	30.09	12.82
3	4	0.65	\$343	4.48	1.03	0.16	4.48	7.85	21.33	7.70
3	4	0.80	\$449	2.84	1.40	0.45	2.84	9.55	13.72	4.80
4	4	0.50	\$228	0.96	0.13	0.13	0.96	10.82	8.73	7.86
4	4	0.65	\$297	0.46			0.46	17.58	13.91	7.21
4	4	0.80	\$385	0.42			0.42	26.42	21.09	10.03
1	5	0.50	\$179	0.06			0.06	5.83	5.83	5.83
1	5	0.65	\$228					2.61	2.61	2.61
1	5	0.80	\$286					2.20	2.20	2.20
2	5	0.50	\$205					16.35	7.70	13.67
2	5	0.65	\$266					19.99	4.21	12.33
2	5	0.80	\$343					28.61	6.25	12.47
3	5	0.50	\$200					5.81	5.81	4.95
3	5	0.65	\$258					2.46	2.46	2.20
3	5	0.80	\$325					1.66	1.66	2.23
4	5	0.50	\$178					2.57	2.57	2.17
4	5	0.65	\$229					1.22	1.22	0.82
4	5	0.80	\$290					0.93	0.93	0.78

Table 3.15: Results: Lead time=1; Trans Matrices 4 & 5

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	6	0.50	\$389	6.49	2.52	0.97	4.82	41.88	40.81	33.81
1	6	0.65	\$497	3.12	1.30	0.34	2.88	21.51	21.02	15.76
1	6	0.80	\$611	1.69	0.44	0.24	1.43	13.64	13.24	9.32
2	6	0.50	\$199	30.79	2.07	0.26	20.33	21.41	19.54	16.96
2	6	0.65	\$290	31.23	5.28	0.16	23.39	14.59	13.69	11.98
2	6	0.80	\$431	25.55	10.57	3.51	20.20	16.69	16.20	16.17
3	6	0.50	\$320	27.12	10.69	1.13	22.47	59.21	22.49	49.56
3	6	0.65	\$452	15.51	9.17	3.13	12.43	26.48	8.01	20.50
3	6	0.80	\$616	10.23	3.97	1.14	7.95	10.53	8.31	9.23
4	6	0.50	\$309	7.86	3.25	0.76	6.68	34.00	32.35	25.56
4	6	0.65	\$412	4.73	1.81	0.41	3.33	14.67	13.94	9.48
4	6	0.80	\$532	1.98	0.87	0.50	1.68	7.06	6.61	6.71
1	7	0.50	\$201	24.62	5.28	2.23	2.61	70.14	6.63	37.33
1	7	0.65	\$299	18.49	6.58	1.99	2.56	38.21	6.41	31.12
1	7	0.80	\$448	17.57	6.15	2.43	1.56	21.07	12.97	25.48
2	7	0.50	\$111	5.76	0.06		0.03	1.83	1.83	1.83
2	7	0.65	\$174	38.64	1.72	0.03	1.47	8.75	8.75	8.75
2	7	0.80	\$285	34.04	13.19	0.51	0.75	15.01	15.01	15.01
3	7	0.50	\$169	45.89	8.96	0.01	1.58	15.45	15.45	15.45
3	7	0.65	\$265	36.50	12.02	3.81	4.18	8.03	8.03	11.02
3	7	0.80	\$416	22.31	8.78	2.08	2.29	9.06	9.06	9.00
4	7	0.50	\$132	10.72	3.62	0.63	0.67	0.86	0.86	0.86
4	7	0.65	\$196	12.30	3.93	0.75	0.84	2.54	2.54	3.08
4	7	0.80	\$300	17.39	5.24	1.46	0.33	6.53	6.53	3.59

Table 3.16: Results: Lead time=1; Trans Matrices 6 & 7

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	1	0.50	\$432	2.07	0.75	0.19	1.90	50.12	50.12	23.77
1	1	0.65	\$554	1.47	0.50	0.09	0.88	33.82	33.82	17.59
1	1	0.80	\$696	0.83	0.21	0.13	0.82	21.35	21.35	12.26
2	1	0.50	\$389	0.64	0.07	0.02	0.60	41.62	41.62	22.52
2	1	0.65	\$517	0.66	0.06	0.01	0.64	61.85	61.85	18.56
2	1	0.80	\$687	0.48	0.03		0.46	124.75	124.75	16.60
3	1	0.50	\$421	4.09	1.29	0.09	4.03	28.12	28.12	15.47
3	1	0.65	\$553	2.68	0.67	0.15	2.64	16.08	16.08	9.82
3	1	0.80	\$715	1.90	0.34	0.28	1.89	8.75	8.75	5.46
4	1	0.50	\$381	1.41	0.36	0.03	1.38	12.61	12.61	5.77
4	2	0.65	\$494	0.71	0.11	0.01	0.68	7.35	7.35	4.15
4	1	0.80	\$635	0.59	0.05		0.58	5.94	5.94	3.18
1	2	0.50	\$412					10.90	18.77	7.84
1	2	0.65	\$544					14.77	28.56	9.14
1	2	0.80	\$713					23.28	61.58	12.00
2	2	0.50	\$413					10.37	21.57	8.09
2	2	0.65	\$546					15.98	32.55	9.31
2	2	0.80	\$716					25.75	68.46	12.31
3	2	0.50	\$417					10.03	22.72	7.57
3	2	0.65	\$551					14.27	26.83	8.77
3	2	0.80	\$722					23.31	73.16	11.60
4	2	0.50	\$412					9.96	19.92	8.14
4	2	0.65	\$544					15.36	28.28	9.54
4	2	0.80	\$713					25.93	56.96	12.89
1	3	0.50	\$414	0.03				4.86	4.86	0.92
1	3	0.65	\$547					0.91	0.91	0.91
1	3	0.80	\$716					1.75	1.75	0.49
2	3	0.50	\$415					2.79	2.79	1.25
2	3	0.65	\$549					2.14	2.14	2.14
2	3	0.80	\$721					6.32	6.32	2.06
3	3	0.50	\$419					3.25	3.25	0.20
3	3	0.65	\$553					0.12	0.12	0.12
3	3	0.80	\$724					2.25	2.25	
4	3	0.50	\$411					3.13	3.13	0.16
4	3	0.65	\$543					0.10	0.10	0.10
4	3	0.80	\$711					2.30	2.30	

Table 3.17: Results: Lead time=2; Trans Matrices 1, 2 & 3

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	4	0.50	\$282	1.21	0.47	0.07	1.21	21.78	19.52	19.99
1	4	0.65	\$359	0.49	0.49	0.05	0.49	17.31	13.92	14.92
1	4	0.80	\$451	0.60	0.05	0.05	0.60	15.36	9.76	10.63
2	4	0.50	\$331	0.62	0.01		0.62	52.88	48.28	32.08
2	4	0.65	\$443	0.91	0.19		0.91	86.65	80.12	24.29
2	4	0.80	\$595	0.39	0.01		0.39	175.56	165.34	29.81
3	4	0.50	\$324	5.24	2.08	0.34	5.24	13.41	48.59	16.85
3	4	0.65	\$428	3.36	0.94	0.28	3.36	10.77	28.04	9.82
3	4	0.80	\$557	2.72	0.60	0.09	2.72	14.68	19.25	6.23
4	4	0.50	\$284	0.55	0.52	0.06	0.55	19.52	7.82	9.20
4	4	0.65	\$369	0.33	0.31		0.33	26.06	9.71	10.27
4	4	0.80	\$477	0.47			0.47	42.86	16.96	10.83
1	5	0.50	\$211					4.61	4.61	4.61
1	5	0.65	\$272					3.60	3.60	3.60
1	5	0.80	\$346					2.31	2.31	2.31
2	5	0.50	\$238					11.81	6.47	11.81
2	5	0.65	\$309					13.37	4.95	10.75
2	5	0.80	\$399					19.21	4.78	11.52
3	5	0.50	\$231					6.41	6.41	4.30
3	5	0.65	\$300					3.05	3.05	2.74
3	5	0.80	\$382					1.80	1.80	1.81
4	5	0.50	\$211					2.95	2.95	1.60
4	5	0.65	\$273					1.47	1.47	1.32
4	5	0.80	\$349					1.25	1.25	0.89

Table 3.18: Results: Lead time=2; Trans Matrices 4 & 5

Prior	P	CR	Cost	Myopic	1 LA	2 LA	OLFC	CEC 1	CEC 2	CEC 3
1	6	0.50	\$517	5.97	2.45	1.10	4.32	70.58	56.78	35.89
1	6	0.65	\$660	3.85	1.25	0.40	2.53	47.76	36.35	30.03
1	6	0.80	\$830	1.99	0.85	0.21	1.74	29.26	21.19	19.69
2	6	0.50	\$259	34.25	2.57	0.29	25.41	35.39	25.46	16.09
2	6	0.65	\$375	35.20	7.68	1.30	28.44	23.66	17.41	13.48
2	6	0.80	\$557	23.67	11.72	3.48	19.29	26.43	24.54	21.69
3	6	0.50	\$425	26.14	13.33	3.12	22.44	84.70	25.22	45.67
3	6	0.65	\$602	16.37	7.53	3.07	13.71	46.59	13.69	26.39
3	6	0.80	\$821	8.23	3.40	1.31	6.40	20.93	18.84	11.25
4	6	0.50	\$407	7.79	3.59	1.22	5.50	55.72	39.63	19.82
4	6	0.65	\$538	4.34	1.45	0.38	3.18	35.45	24.19	13.75
4	6	0.80	\$697	2.35	0.79	0.35	2.03	19.31	14.81	7.84
1	7	0.50	\$221	25.34	6.92	1.34	1.56	57.01	5.30	37.76
1	7	0.65	\$327	21.69	6.89	2.44	1.28	28.66	3.70	28.66
1	7	0.80	\$489	14.89	7.42	1.93	1.11	15.53	14.97	22.90
2	7	0.50	\$125	9.79	0.08		0.07	2.35	2.35	2.35
2	7	0.65	\$193	30.24	0.51	0.12	0.23	3.40	3.40	3.40
2	7	0.80	\$306	35.26	9.80	1.42	1.61	18.72	18.72	18.72
3	7	0.50	\$185	38.37	6.67	0.76	3.09	12.47	12.47	18.28
3	7	0.65	\$287	30.74	13.63	2.44	2.77	4.69	4.69	7.14
3	7	0.80	\$445	18.36	9.57	2.56	1.38	10.77	10.77	9.79
4	7	0.50	\$146	8.73	2.49	0.25	0.28	0.56	0.56	2.75
4	7	0.65	\$215	15.01	2.19	0.12	0.20	0.58	0.58	0.68
4	7	0.80	\$324	17.11	6.84	0.54	0.67	9.23	9.23	4.85

Table 3.19: Results: Lead time=2; Trans Matrices 6 & 7

CHAPTER 4

BOUNDS AND APPROXIMATIONS FOR ADAPTIVE INVENTORY CONTROL

4.1 Introduction

In practice, product demand is typically highly uncertain and subject to change, and the demand process is only partially observed because the probability distribution of demand in a given period is not known with certainty. Moreover, this demand distribution may change over time according to some random process. Because optimal solutions for this problem are available only for small problems, managers must estimate the unknown demand distribution. This estimate is commonly treated as if it were known with certainty, and then some type of certainty equivalence control strategy is used to generate a control policy.

Effective approximations to the optimal cost function are known for the stationary version of this problem, and we extend these procedures to demonstrate that similar techniques can be used when the demand distribution is *non-stationary*. We consider an infinite horizon problem with positive lead time, full backordering, and linear holding and backorder costs. We model the problem as a composite-state, partially observed Markov decision process. The demand distribution is assumed to belong to a known and discrete set of candidate distributions, and a known, discrete *a priori* distribution is

assumed for this set which is also used along with subsequent observations to compute a *posterior* distribution.

Optimal solutions to this problem are typically intractable because the state space for the prior is uncountably infinite. Therefore, we first develop a grid approximation that is asymptotically exact for this problem. We then show how to calculate both an upper bound and a lower bound for the optimal cost function, and we compare these bounds to the functional approximation. For a stationary problem, the lower and upper bounds are tight at the extreme points. However, for the non-stationary problem, there is a gap between the upper and lower bounds at the extreme points. Furthermore, the initial bounds are not as tight away from the extreme points as in the stationary problem due to the added uncertainty generated by the non-stationary aspect of the demand process. However, both of these bounds can be iteratively tightened. Additionally, we describe details for myopic and look-ahead policies for this problem.

In Section 4.2 we review key literature, and in Section 4.3 we review the basic control model. In Section 4.4 we discuss both the grid approximation as well as upper and lower bounds and the iterative tightening of those bounds, and the myopic and look-ahead policies. In Section 4.5 we provide an example of our technique, and finally, we provide conclusions and recommendations for future research in Section 4.6.

4.2 Literature Review

Non-stationary, partial information problems are, by their very nature, difficult to solve. Therefore, the literature for this class of problems is relatively sparse. Most traditional inventory control models assume either stationary demand or complete information about the underlying demand state or both. Hadley and Whitin [13] solve a stochastic non-stationary problem. They determine a final inventory level for a known obsolescence date with Poisson demand. Karlin [17] studies non-stationary problems and presents several qualitative characteristics of the optimal inventory level. For example, he notes that the critical level will decrease if the demand densities are decreasing over consecutive periods but may or may not increase if the demand densities are increasing over consecutive periods. In this same work, he assumes that the demand is stationary but that the distribution is only known partially, that is, the distribution has an unknown parameter with a known *a priori* distribution. He presents several results about the characteristics of optimal policies for the exponential (including gamma, Poisson, and negative binomial) and range families, both of which have a single sufficient statistic for the unknown parameter. Veinott [42] extended Karlin's work by showing that when one density is a translate of another, it is possible to relax the requirement for stochastic ordering to show the relationship between two sets of critical numbers. He also shows how to change a zero lead time problem with no backlogging to one with backlogging.

More recently, Song and Zipkin [37] present a general modeling framework for non-stationary demand. They demonstrate that optimal policies behave very similarly to

the newsboy problem in their dependence on unit cost, holding cost, and penalty cost, and they present two algorithms to solve the linear cost problem. In a later work, Song and Zipkin [38] examine a problem with obsolescence. But they always assume that the current demand distribution is always known with certainty. They model the demand process as a continuous time, non-increasing discrete state Markov chain. They demonstrate the performance of a blind policy which forecasts demand only over the lead time and assumes that the demand distribution does not change over the lead time. They also examine a limited look-ahead policy that accounts for demand changes over the lead time but not beyond. Stephen Graves [12] presents an ARIMA based model to compute an adaptive base stock policy for non-stationary demand. Again, this problem can be assumed to have complete information because the distribution of the demand in a period is fully determined by the ARIMA model and the observed demand in the previous period. Gavirneni and Tayur [11] present an efficient procedure to compute the state dependent policy for a non-stationary problem. They do not calculate the actual cost of the policy but rather compute the derivative of the cost function to determine the optimal base stock level for each state. However, the optimal cost can be calculated using other methods once the optimal policy is known.

Partial information problems are more difficult to solve than the complete information problems addressed by the papers above. Bayesian methods or other adaptive procedures can be used to solve these problems, but they often require a great deal more computational effort. Herbert Scarf [33] was a pioneer in the use of Bayesian techniques

for inventory control when he studied an inventory problem with linear holding and penalty costs. The demand density was described by a density $\phi(\xi, \omega)$ with unknown parameter ω . The parameter ω , however, has a known *a priori* distribution. The demand density was restricted to the exponential family because a single sufficient statistic, S , can summarize all prior demand information. For the exponential family, the sufficient statistic, S , is the average demand over the previous periods. Scarf shows that the optimal stock level can be calculated recursively with knowledge of the current inventory level, x , and the sufficient statistic, S . In a subsequent paper [34], he shows that in some instances, the critical level can be determined as a function of one variable if the demand distribution is in the gamma family.

Katy Azoury [4] demonstrated that other problems can also be solved using a one-dimensional state space when there is a natural conjugate family. If the demand distribution has a fixed dimension sufficient statistic, and the distribution of the unknown parameter comes from a conjugate family, then the posterior distribution for the unknown parameter is also in that same family. She then presents two conditions that show when the model can be reduced to a one-dimensional state space.

Lovejoy [25] uses a myopic parameter adaptive technique and a simple inventory policy based upon a critical fractile. Specifically, he uses exponential smoothing and Bayesian updating of parameter estimates. He also determines bounds on the value loss relative to optimal costs when using his policy. Lovejoy [28] provides critical insights into the stationary, partial information problem that address bounds for certain suboptimal

policies, and he shows how it is possible to iteratively tighten these bounds. He also addresses a problem in which some parameters may be non-stationary, although the underlying demand distribution remains stationary. Lovejoy models the problem as a composite-state partially observed Markov decision process (POMDP) in which the inventory level is observed completely while the demand distribution is observed partially. This work extends this approach to the problem in which the demand distribution is non-stationary.

Problems with non-stationary demand and partial information present an even more complex situation. Little direct work has been done in this area, but there is great potential impact for research on this class of problems. One paper that considers a problem in this class is by Kurawarwala and Matsuo [19]. They present a growth model to estimate the parameters of a demand process over its entire life cycle. In their base case, production decisions are made at the beginning of the problem for the entire life cycle. They present a technique to initially estimate the parameters of their forecasting model. However, they do not thoroughly address the issue of revising these estimates using new observations.

Treharne and Sox [40] examine an inventory control problem in which the demand process is partially observed and non-stationary. Although managers often use certainty equivalent control (CEC) policies to solve such a problem, Treharne and Sox demonstrate that there exist other efficient and often superior suboptimal control policies to solve this problem. They model the problem as a composite-state, partially observed Markov

decision process over a finite horizon. They show that consideration of the uncertainty of the demand distribution, even using a suboptimal policy over a finite horizon often leads to substantial improvements over CEC policies.

We model this inventory control problem as a composite-state, partially observed Markov decision process (CPOMDP), see Section 4.3. In a CPOMDP, some states are fully observed while others are only partially observed. Partially observed Markov decision processes have been well studied in the past, but with few applications in inventory and production control. However, it presents a body of theory with some direct application to our problem. We, therefore, present several key references for partially observed Markov decision processes. White and White [45] give an excellent review of standard Markov decision processes. They include a short section on extensions of MDPs including that of partially observed MDPs. Monahan [29] and Lovejoy [27] present two good surveys of POMDP. Smallwood and Sondik [35] show that for a finite horizon POMDP maximization problem the objective function is piece-wise linear and convex in the current state probabilities. They then provide a dynamic programming algorithm that uses this property to efficiently solve the problem. Sondik [36] later presents a policy iteration technique for a discounted, infinite horizon problem. White and Sherer [43] present three algorithms for the POMDP that are faster than that of Smallwood and Sondik. However, the problem sizes are very limited in their study. Lovejoy [24] [26] presents monotonicity results and provides feasible bounds for certain POMDP problems. White

and Sherer [44] propose a heuristic in which only the M most recent observations and actions are used to make the current decision. They also introduce the concept of ergodicity and discuss how this factor indicates the potential value of historical information.

4.3 Optimal Control Model and Insights

In this section, we summarize the optimal control model and discuss insights gained from its structure. Demand each period arises from one of a finite collection of N probability distributions. The decision maker may not know which of the distributions generates the demand in a given period. Furthermore, the distribution randomly changes from one period to the next according to a known transition probability matrix. In this chapter, we assume that the planning horizon is infinite with discounted costs, stockouts are always backordered, the order lead time is either zero or a known positive constant, and there is no fixed order cost. We define in Table 4.1 the notation that will be used to describe the model.

Although we specify linear holding and stockout costs in this basic model, many of the results hold for convex costs. The demand state process $\{d_t\}$ is not known with certainty and is referred to as the *core process*. It is modeled as a finite state Markov chain, $d_t \in \{1, \dots, N\}$, governed by the transition matrix P . If $d_t = j$, the probability distribution for demand in t is r_j , i.e., $P[w_t = k | d_t = j] = r_{jk}$ for $k = 0, 1, \dots, M$. We assume all cost factors and transition equations do not change over time; therefore, the optimal policy, δ , will also be stationary. The core process is partially observed through

N	=	number of distributions,
M	=	maximum demand,
S	=	stock level,
h	=	linear holding cost,
p	=	linear stockout penalty cost,
c	=	linear procurement cost,
β	=	discount rate,
l	=	lead time,
d_t	=	demand distribution in effect for period t ,
p_{ij}	=	$P[d_{t+1} = j d_t = i]$,
P	=	(p_{ij}) , the transition probability matrix for d_t ,
w_t	=	demand realization in period t ,
$w_{t,l}$	=	lead time demand realization from period $t \rightarrow t + l$,
r_{jk}	=	$P[w_t = k d_t = j]$,
r_{jkl}	=	$P[w_{t,l} = k d_t = j]$,
I_t	=	vector of information available up to period t ,
π_{it}	=	$P[d_t = i I_t]$,
π_t	=	(π_{it}) ,
x_t	=	inventory level at the beginning of period t ,
u_t	=	inventory position at the beginning of period t ,
a_t	=	order quantity in period t ,
$\bar{a}_{t,l}$	=	vector of outstanding orders,
δ	=	stationary policy.

Table 4.1: Notation

the *measurement process* $\{w_t\}$, and the vector of information available at time t is $I_t = (w_0, w_1, \dots, w_{t-1})$. Also known are π_t and u_t where π_0 is the initial prior distribution that is externally specified. Rhenius [32] has shown that the prior distribution, π_t , is a sufficient statistic for I_t which implies that the problem can be viewed as a Markov decision process on the state space (u_t, π_t) . The vector π_t characterizes the current belief of the distribution of d_t given all prior observations of the information process $\{w_t\}$. The system dynamics for our CPOMDP model are described in Figure 4.1. The

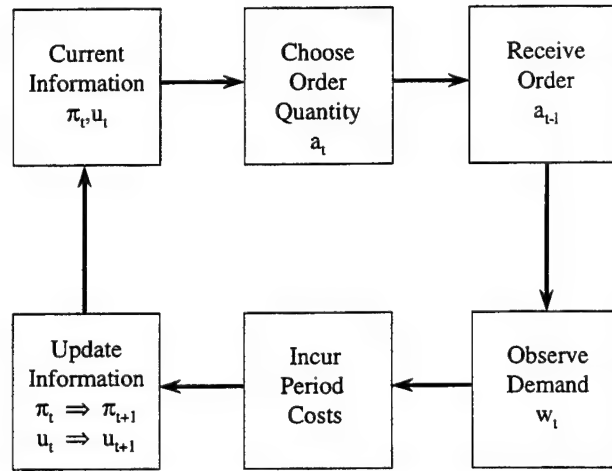


Figure 4.1: Control Process

control process begins with an information vector, I_t , which includes a prior distribution, π_t , for period t , and then selects an order quantity, a_t . Because any optimal policy, δ^* , is stationary, the order quantity at time t is a function of the current prior, π_t , and inventory position, u_t . For a positive lead time, an order placed l periods in the past is received. The demand, w_t , then occurs, and the inventory costs for period t are incurred. The time index advances to $t+1$, and the core process advances from d_t to d_{t+1} according to the transition matrix P . The new prior distribution π_{t+1} is computed using π_t, w_t and P . The transition equation for the prior is

$$\pi_{i,t+1} = T_i(\pi_t | w_t = k) = \frac{\sum_{j=1}^N \pi_{jt} r_{jk} p_{ji}}{\sum_{j=1}^N \pi_{jt} r_{jk}}. \quad (4.1)$$

The problem is formulated in terms of the *inventory position*, which is the net inventory level (on-hand less backorders) plus inventory on-order at the beginning of t , $u_t = x_t +$

$\sum_{n=1}^l a_{t-n}$. The inventory position follows the transition equation $u_{t+1} = u_t + a_t - w_t$.

For an inventory level v , define the single period expected inventory cost function as

$$G_t(v|\pi_t, l) = E_{w_{t,l}|\pi_t}[p \max\{0, w_{t,l} - v\} + h \max\{0, v - w_{t,l}\}], \quad (4.2)$$

where $w_{t,l} = \sum_{n=0}^l w_{t+n}$. This cost function accounts for the positive lead time in computing the expected inventory costs for the inventory level at time $t+l$. The optimal infinite horizon cost-to-go function satisfies the well known Bellman's equation [5]:

$$J^*(u, \pi) = -cu + \min_{S \geq u} \{cS + G(S|\pi, l) + \beta E_{w|\pi}[J^*(S - w, T(\pi|w))]\}. \quad (4.3)$$

In this equation, w is the lead time demand given the current prior, π . A state dependent base stock control policy is optimal for this problem. Treharne and Sox [40] demonstrate that a base stock policy is optimal for the finite horizon problem because the cost function is a piecewise-linear and convex (pwl-convex) function of the inventory level/position. A simple extension of this proof shows that a base stock policy is optimal for the infinite horizon problem also. This form suggests that $J^*(u, \pi)$ might be solved using a standard policy iteration or value iteration procedure. Although the state space on u is a finite set, the state space on π is uncountably infinite. Therefore, some approximation procedure is necessary for most problems.

This problem has some very useful properties. Smallwood and Sondik [35] show that for a finite horizon problem the optimal cost function is a piecewise-linear and

concave (pwl-cc) function of π . Although their problem was for a partially observed Markov decision process, the property extends to our problem which is a composite-state, partially observed Markov decision process. For a given inventory position, u , the cost function, $J^*(u, \pi)$, for the CPOMDP is also a pwl-cc of π when the horizon is finite. However, in our CPOMDP problem, the horizon is infinite. We conjecture that for a given inventory position, u , the optimal cost-to-go function will be concave in π , but not necessarily piecewise-linear. The cost function is easier to compute if it is piecewise-linear. Sondik [36] shows that the cost function for a partially observable Markov decision process is piecewise-linear if the optimal policy belongs to the class of *finitely transient* policies. Not all policies have this property, and additionally, there are no simple rules for determining if a policy is finitely transient. Because the cost function for our problem is concave in π , we can approximate the optimal policy by a finite number of subgradients to the cost function. We will show in Section 4.4 how these subgradients can be computed.

4.4 Suboptimal Control

4.4.1 Grid Approximation

To evaluate these policies, we would ideally compare the suboptimal policies to known optimal results. Unfortunately, because the state space, Π , is uncountably infinite, we cannot do so. Therefore, we have chosen to approximate the optimal cost function using a grid approximation method. We do so by discretizing the state space on Π .

Using $G + 1$ grid points for π_i such that

$$\hat{\Pi} = \left\{ \pi^k \in R^n : \sum_{i=1}^N \pi_i^k = 1, \pi_i^k \in \left\{ 0, \frac{1}{G}, \frac{2}{G}, \dots, 1 \right\} \right\}. \quad (4.4)$$

We index the j^{th} element of each set as π^j . It is necessary to calculate the probability of transitioning from one allowable discrete state in $\hat{\Pi}$ to another discrete state in $\hat{\Pi}$. The updated prior, $\pi_{t+1} = T_i(\pi_t | w_t = k)$, will most likely not be a feasible value of π , that is, it will not be a member of $\hat{\Pi}$. Thus, we find the nearest feasible grid points in $\hat{\Pi}$ that contain π_{t+1} . We then calculate the transition probability that the updated prior will move to a particular feasible value of π in the grid. The barycentric coordinates of the updated prior with respect to these nearest feasible grid points represent the transition probabilities to these points. Lovejoy [26] provides details and demonstrates an efficient way to calculate these barycentric coordinates using a Freudenthal Triangulation [9] technique and a transformation of the state space, $\hat{\Pi}$. This triangulation technique allows any point in the state space to be quickly triangulated. However, if the triangulation is done only in $\hat{\Pi}$, then the resulting barycentric coordinates may not be at feasible vertices in $\hat{\Pi}$. We briefly review these procedures in the context of our problem. Given an updated prior, π_{t+1} , we must estimate the probabilities that the updated state in $\hat{\Pi}$ will be at a finite number of feasible vertices.

Calculation of updated state space in $\hat{\Pi}$:

Define an $N \times N$ matrix:

$$B = \frac{1}{G} \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \dots & 1 & -1 \\ 0 & 0 & \dots & \dots & \dots & \dots & 1 \end{pmatrix}$$

1. Choose a grid spacing with G intervals.
2. For any posterior probability, π_{t+1} , calculate the coordinates of x in transformed state space where $x = B^{-1}\pi_{t+1}$. Note that B^{-1} has a simple form; it is an upper triangle matrix with all elements equal to G .

$$B^{-1} = \begin{pmatrix} G & \dots & \dots & \dots & \dots & \dots & G \\ 0 & G & \dots & \dots & \dots & \dots & G \\ 0 & 0 & G & \dots & \dots & \dots & G \\ \vdots & \dots & \dots & \dots & G & G & G \\ 0 & 0 & \dots & \dots & \dots & G & G \\ 0 & 0 & \dots & \dots & \dots & \dots & G \end{pmatrix}$$

3. Determine vertices that contain x using Freudenthal Triangulation. The base vertex, v_1 , is equal to the largest integer vector less than or equal to x . Calculate the

direction, d , from the base vertex, v_1 , to x , $d = x - v_1$. Order the components of d in descending order with p being the permutation of the integers $\{1, \dots, n\}$ that indicates the ordering ($d_{p1} \geq d_{p2} \geq \dots \geq d_{pn}$). Define e_j as the j^{th} unit vector. The remainder of the vertices that form the triangulation are

$$v_2 = v_1 + e_{p1}$$

$$v_3 = v_2 + e_{p2}$$

$$v_4 = v_3 + e_{p3}$$

$$\vdots$$

$$v_{k+1} = v_k + e_{pk} \quad k \leq n.$$

4. Determine the barycentric coordinates λ_i using Freudenthal Triangulation.

$$\lambda_{n+1} = d_{pn}$$

$$\lambda_n = d_{p(n-1)} - d_{pn}$$

$$\lambda_{n-1} = d_{p(n-2)} - d_{p(n-1)}$$

$$\vdots$$

$$\lambda_2 = d_{p1} - d_{p2}$$

$$\lambda_1 = 1 - \sum_{i=2}^{n+1} \lambda_i$$

5. Transform the vertices with positive λ_i back to state space $\hat{\Pi}$.

$$\pi^1 = B * v^1$$

$$\pi^2 = B * v^2$$

$$\vdots$$

$$\pi^{n+1} = B * v^{n+1}$$

The transition probability from π_{t+1} to $\pi^j = \lambda_j$.

Consider the following example. Assume that $N = 4$ distributions, and that $G = 10$ which is spacing = .01. Determine the vertices and transition probabilities if the updated prior, $\pi_{t+1} = [.48, .28, .13, .11]$. First, $x = B^{-1}\pi_{t+1} = [10, 5.2, 2.4, 1.1]$.

1. Determine the vertices that triangulate x .

$$(a) \ v_1 = [10, 5, 2, 1] \quad d = [0, .2, .4, .1] \quad p = [3, 2, 4, 1]$$

$$(b) \ v_2 = [10, 5, 2, 1] + [0, 0, 1, 0] = [10, 5, 3, 1]$$

$$(c) \ v_3 = [10, 5, 3, 1] + [0, 1, 0, 0] = [10, 6, 3, 1]$$

$$(d) \ v_4 = [10, 6, 3, 1] + [0, 0, 0, 1] = [10, 6, 3, 2]$$

$$(e) \ v_5 = [10, 6, 3, 2] + [1, 0, 0, 0] = [11, 6, 3, 2]$$

2. Determine the barycentric coordinates for x relative to vertices.

$$(a) \lambda_5 = 0.0$$

$$(b) \lambda_4 = 0.1 - 0.0 = 0.1$$

$$(c) \lambda_3 = 0.2 - 0.1 = 0.1$$

$$(d) \lambda_2 = 0.4 - 0.2 = 0.2$$

$$(e) \lambda_1 = 1 - 0.2 - 0.1 - 0.1 - 0.0 = 0.6$$

3. Vertices transformed to $\hat{\Pi}$.

$$(a) \pi^1 = B * v^1 = [0.5, 0.3, 0.1, 0.1]$$

$$(b) \pi^2 = B * v^2 = [0.5, 0.2, 0.2, 0.1]$$

$$(c) \pi^3 = B * v^3 = [0.4, 0.3, 0.2, 0.1]$$

$$(d) \pi^4 = B * v^4 = [0.4, 0.3, 0.1, 0.2]$$

$$(e) \pi^5 = B * v^5 = [0.5, 0.3, 0.1, 0.2]$$

First, note that π^5 is infeasible. However, $\lambda_5 = 0.0$, and there is no transition probability to π^5 . Therefore, if the updated prior is $\pi_{t+1} = [.48, .28.13, .11]$, we estimate a transition probability of 0.6, 0.2, 0.1, and 0.1 to feasible vertices π^1, π^2, π^3 , and π^4 respectively. The procedures to find the feasible vertices and respective transition probabilities are straightforward to implement. B^{-1} is simple to calculate, and both the vertices and barycentric coordinates are computationally easy to determine.

The linear approximation allows us to represent every feasible transition in the state space. Using this grid approximation and linear interpolation, the cost-to-go function is

$$J_G(u, \pi^k) = \min_{S \geq 0} \left\{ c \max(0, S - u) + G(\max(u, S) | \pi^k, l) \right. \\ \left. + \beta E_{w | \pi^k} \left[\sum_{j=1}^{N(G+1)} \lambda_j(T(\pi^k | w)) J_G(\max(u, S) - w, \pi^j) \right] \right\} \quad (4.5)$$

where $\lambda_j(T(\pi^k | w))$ are the coordinates of the Freudenthal triangulation of $T(\pi^k | w)$ over the grid $\hat{\Pi}$. Therefore, for each grid point π^k and each possible value of w , we must compute the Freudenthal Triangulation of $T(\pi^k | w)$. Finally, the cost for any value of π can be computed by

$$J_G(u, \pi) = \sum_{i=1}^{G+1} \lambda_i J_G(u, B * v_i) \quad (4.6)$$

Lovejoy [26] shows that $J_G(u, \pi)$ is also a lower bound to the optimal cost function.

4.4.2 Upper Bound

We next extend the upper bounding procedure described in Lovejoy [28] for this POMDP problem. In the stationary problem studied in Lovejoy [28], if the partially observed state becomes known with certainty, the problem is then *fully observed* and is solvable by conventional dynamic programming algorithms. Van Hee [41] exploits this fact to show how upper and lower bounds can be developed by solving *fully observed*

problems. This characteristic of the stationary problem also means that Lovejoy's upper bound is tight at the extreme points. In the case of a non-stationary problem, even if we know the current state with certainty, we will not know the future states with certainty. Therefore, it is not clear how Lovejoy's procedure can be effectively extended to the case of a non-stationary process.

In order to develop our upper bounding procedure, we first consider the class of static base stock policies. A static base stock policy chooses a base stock level S given an initial prior, π , and uses that same base stock level in all future periods. Such a policy is clearly suboptimal, and as such it provides an upper bound for the optimal cost function. For an initial inventory position u , an initial prior distribution π , and a base stock level S , the infinite horizon discounted cost is given by

$$J(u, \pi, S) = c \max(0, S - u) + G(\max(u, S) | \pi, l) + \beta E_{w | \pi} [J(\max(u, S) - w, T(\pi | w), S)]. \quad (4.7)$$

Because this policy does not use information about the future values of π , Equation (4.7) can be re-written as

$$J(u, \pi, S) = c \max(0, S - u) + \sum_{i=1}^N \pi_i G(\max(u, S) | e_i, l) + \beta \sum_{i=1}^N \sum_{j=1}^N \pi_i P_{ij} E_{w | e_i} [J(\max(u, S) - w, e_j, S)] \quad (4.8)$$

which is readily computed as the solution of a system of linear equations for a fixed value of π . Also, note that $J(u, \pi, S)$ is linear in π . The infinite horizon discounted cost of the optimal static base stock policy is

$$J_S(u, \pi) = \min_{S \geq 0} J(u, \pi, S) \quad (4.9)$$

which is piecewise linear and concave in π because $J(u, \pi, S)$ is a linear function of π for each S . Define $S_i = \operatorname{argmin}_{S \geq 0} J(u, e_i, S)$ which is the optimal static base stock level given that demand distribution i is known to be the initial demand distribution, i.e., $J_S(u, e_i) = J(u, e_i, S_i)$. For $j = 1, \dots, N$, define the vector $\psi^j(u)$ as

$$\begin{aligned} \psi_i^j(u) = J(u, e_i, S_j) = & c \max(0, S_j - u) + G(\max(u, S_j) | e_i, l) \\ & + \beta \sum_{k=1}^N P_{ik} E_{w|e_i} [J(\max(u, S_j) - w, e_k, S_j)]. \end{aligned} \quad (4.10)$$

From Equations (4.8) and (4.9), we see that $\psi^j(u)$ is a subgradient of J_S at the point (u, e_j) because $J_S(u, e_j) = J(u, e_j, S_j)$ and

$$\begin{aligned} J(u, \pi, S_j) &= \sum_{i=1}^N \pi_i \left[c \max(0, S_j - u) + G(\max(u, S_j) | e_i, l) \right. \\ &\quad \left. + \beta \sum_{k=1}^N P_{ik} E_{w|e_i} [J(\max(u, S_j) - w, e_k, S_j)] \right] \\ &= \sum_{i=1}^N \pi_i \psi_i^j(u). \end{aligned} \quad (4.11)$$

Furthermore, we have that

$$J(u, \pi, S_j) \geq J_S(u, \pi) \geq J^*(u, \pi) \text{ for all } j = 1, \dots, N.$$

Therefore,

$$\bar{J}(u, \pi) = \min_{j=1, \dots, N} \sum_{i=1}^N \pi_i \psi_i^j(u) \geq J_S(u, \pi, S) \geq J^*(u, \pi). \quad (4.12)$$

The function $\bar{J}(u, \pi)$ is also a piecewise linear function of π . Because it is based on a set of static base stock policies, it may admittedly be a weak bound in some cases. Therefore, we would like to improve this bound at values of π where it is not tight. Such a refinement can be accomplished, as in Lovejoy [28], by using one or more dynamic programming iterations at the chosen point(s). Define $l(t) = \operatorname{argmin}_j \sum_{i=1}^N \pi_i \psi_i^j(u)$, and the one-step DP recursion is

$$\begin{aligned} \bar{J}^1(u, \pi) = H\bar{J}(u, \pi) = \min_{S \geq 0} & \left\{ c \max(0, S - u) + G(\max(u, S) | \pi, l) \right. \\ & \left. + \beta E_{w|\pi} [\bar{J}(\max(u, S) - w, T(\pi|w))] \right\}. \end{aligned} \quad (4.13)$$

As demonstrated by Lovejoy [28],

$$\bar{J}(u, \pi) \geq \bar{J}^1(u, \pi) \geq J^*(u, \pi).$$

Note that the expected future cost in Equation (4.13) can be expressed as

$$\sum_{i=1}^N \pi_i \sum_{k=0}^M r_{ik} \sum_{j=1}^N \frac{\sum_{n=1}^N \pi_n r_{nk} p_{nj}}{\sum_{n=1}^N \pi_n r_{nk}} \cdot \psi^{l(T(\pi|k))}(\max(u, S) - k)$$

$$= \sum_{k=0}^M \sum_{j=1}^N \sum_{n=1}^N \pi_n r_{nk} P_{nj} \cdot \psi^{l(T(\pi|k))}(\max(u, S) - k)$$

$$= \sum_{i=1}^N \pi_i \sum_{k=0}^M r_{ik} \sum_{j=1}^N P_{ij} \cdot \psi^{l(T(\pi|k))}(\max(u, S) - k)$$

which is piecewise linear and concave in π . Therefore, $\bar{J}^1(u, \pi)$ is piecewise linear and concave in π . Let $j = N + 1$ and define

$$\begin{aligned} \psi_i^{N+1}(u) &= c \max(0, S_{N+1} - u) + G(\max(u, S_{N+1})|e_i, l) \\ &\quad + \beta \sum_{k=0}^M \sum_{j=1}^N r_{ik} P_{ij} \psi_j^{l(T(\pi|k))}(\max(u, S_{N+1} - k) \end{aligned} \quad (4.14)$$

for some π . Then, $\psi^{N+1}(u)$ given by Equation (4.14) is a subgradient of \bar{J}^1 at (u, π) .

Furthermore, $\bar{J}(u, \pi) = \min_{j=1, \dots, N+1} \sum_{i=1}^N \pi_i \psi_i^j(u) \geq \bar{J}^1(u, \pi) \geq J^*(u, \pi)$. Note that

when $P = I$, as in Lovejoy [28], Equation 4.14 simplifies to

$$\psi_i^{N+1}(u) = c \max(0, S - u) + G(\max(u, S) | e_i, l) + \beta \sum_{k=0}^M r_{ik} \psi_i^{l(T(\pi|k))}(\max(u, S - k)). \quad (4.15)$$

Thus, our procedure is to first initialize the upper bound function $\bar{J}(u, \pi)$ using Equation (4.10) for $j = 1, \dots, N$. Then for some selected values of π , say π^1, \dots, π^K , Equation (4.14) is used to generate additional subgradients $\psi^{N+1}, \dots, \psi^{N+K}$. The upper bounding function is given by $\bar{J}(u, \pi) = \min_{j=1, \dots, N+K} \sum_{i=1}^N \pi_i \psi_i^j(u)$. This process can be repeated to add additional subgradients to the definition of $\bar{J}(u, \pi)$.

It should be noted that the upper and lower bounds are tight at the extreme point for the stationary problem. In the non-stationary, there is a gap between these bounds. Figure 4.2 shows the gap between these bounds. The improvement steps discussed tighten these bounds.

4.4.3 Lower Bound

We now present a lower bounding procedure.

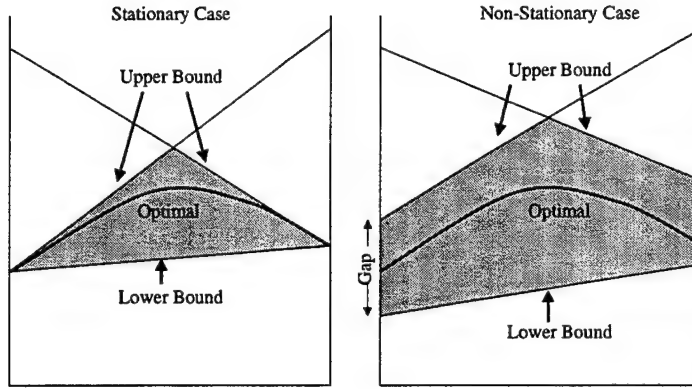


Figure 4.2: Comparison of Stationary and Non-Stationary Bounds

1. Initialize the lower bound by solving the fully observed problem at the extreme points of Π , e_i , and let $\hat{J}^o(u, e_i) = \text{value at } e_i$, where

$$\hat{J}^o(u, e_i) = \min_{S_i \geq u} \left\{ c(S_i - u) + G(S_i | d = i, l) + \beta \sum_{j=1}^N E_{w|d=i} [P_{ij} \hat{J}^o(S_i - w, e_j)] \right\} \quad (4.16)$$

2. Estimate J at a fixed set of grid points using a convex combination of $\{\hat{J}^o(u, e_i)\}$:

$$\hat{J}^o(u, \pi^k) = \sum_{i=1}^N \pi_i^k \hat{J}^o(u, e_i).$$

3. Use a sequence of DP recursions to update values at the fixed set of grid points so that

$$\begin{aligned} \hat{J}^n(u, \pi^k) = \min_{S \geq 0} & \left\{ c \max(0, S - u) + G(\max(u, S) | \pi^k, l) \right. \\ & \left. + \beta E_{w | \pi^k} \left[\sum_{j=1}^{N(G+1)} \lambda_j(T(\pi^k | w)) \hat{J}^{n-1}(\max(u, S) - w, \pi^j) \right] \right\}, \text{ for } n = 1, 2, \dots, \end{aligned}$$

where $\{\lambda_j(T(\pi^k | w))\}$ are the coordinates of the Freudenthal triangulation of $T(\pi^k | w)$ over grid G .

We make the following two conjectures which assume that the same $\hat{\Pi}$ is used for both J_G and $\hat{J}^n \forall n$.

Conjecture 1. $\hat{J}^n(u, \pi^k) \leq J_G(u, \pi^k) \quad \forall n, u, \text{ and } \pi^k$.

Conjecture 2. $\lim_{n \rightarrow \infty} \hat{J}^n(u, \pi^k) = J_G(u, \pi^k) \quad \forall u, \text{ and } \pi^k$.

4.4.4 Myopic and Limited Look-Ahead Policies (Infinite Horizon)

Define

$$\bar{J}^{L0}(u, \pi) = \min_{S \geq u} \{ c(S - u) + G(S | \pi, l) \} \quad (4.17)$$

as the myopic cost function which can be re-written as

$$\bar{J}^{L0}(u, \pi) = \min_{S \geq u} \left\{ \sum_{i=1}^N \pi_i [c(S - u) + G(S | e_i, l)] \right\}. \quad (4.18)$$

Therefore, $\bar{J}^{L0}(u, \pi)$ is piecewise linear and concave in π . Let

$$\alpha_i^j(u) = c \max(0, S_j - u) + G(\max(u, S_j)|e_i, l),$$

then

$$\bar{J}^{L0}(u, \pi) = \min_j \sum_{i=1}^N \pi_i \alpha_i^j(u). \quad (4.19)$$

Because $\sum_{i=1}^N \pi_i \alpha_i^j(u)$ is a linear function of π , $\bar{J}^{L0}(u, \pi)$ is piecewise linear and concave in π .

The one-period look-ahead cost function is

$$\bar{J}^{L1}(u, \pi) = \min_{S \geq u} \left\{ c(S - u) + G(S|\pi, l) + E_{w|\pi} \left[\bar{J}^{L0}(S - w, T(\pi|w)) \right] \right\} \quad (4.20)$$

which can be re-written as

$$\bar{J}^{L1}(u, \pi) = \min_{S \geq u} \left\{ \sum_{i=1}^N \pi_i \left[c(S - u) + G(S|e_i, l) + E_{w|e_i} \left[\sum_{j=1}^N p_{ij} \alpha_j^{l(T(\pi|w))} (S - w) \right] \right] \right\} \quad (4.21)$$

where $l(T(\pi|w)) = \operatorname{argmin}_j \sum_{i=1}^N T_i(\pi|w) \alpha_i^j(S - w)$. Therefore, $\bar{J}^{L1}(u, \pi)$ is also piecewise linear and concave in π . Cheng's [7] linear support algorithm can be used to identify all the subgradients needed to define $\bar{J}^{L1}(u, \pi) = \min_j \sum_{i=1}^N \pi_i \alpha_i^j(u)$.

1. Calculate the subgradient $\gamma^j(u)$ at each extreme point, e_j , of Π and store in the set $G_t(u)$.
2. For each $\gamma^j(u) \in G_t(u)$, identify an extreme point, π^{jk} , of

$$R(\gamma^j(u)) = \{\pi | \pi^T \gamma^j(u) \leq \pi^T \gamma^l(u) \forall l \neq j\}$$

by solving for some $k \neq j$

$$\min \pi^T \gamma^k(u)$$

$$s.t. \pi^T (\gamma^j(u) - \gamma^l(u)) \leq 0 \quad \forall l \neq j$$

$$\sum \pi_i = 1$$

$$\pi_i \geq 0 \quad \forall i$$

3. Define $\bar{J}(u, \pi) = \min_j \pi^T \gamma^j(u)$, and

$$\bar{J}^{L1}(u, \pi) = \min_{S \geq u} \left\{ c(S - u) + G(S | \pi, l) + E_{w | \pi} [\bar{J}^{L0}(S - w | T(\pi | w))] \right\}.$$

If $\bar{J}(u, \pi^{jk}) > \bar{J}^{L1}(u, \pi^{jk})$, then compute the subgradient, $\gamma(u)$, of \bar{J}^{L1} at π^{jk} , and add to $G_t(u)$, and go to 2. Otherwise, next k or next j or end.

Dist.	Initial		Truncated
	Mean	Std. Dev.	Mean
1	2.00	1.5	2.00
2	14.00	4.0	13.45

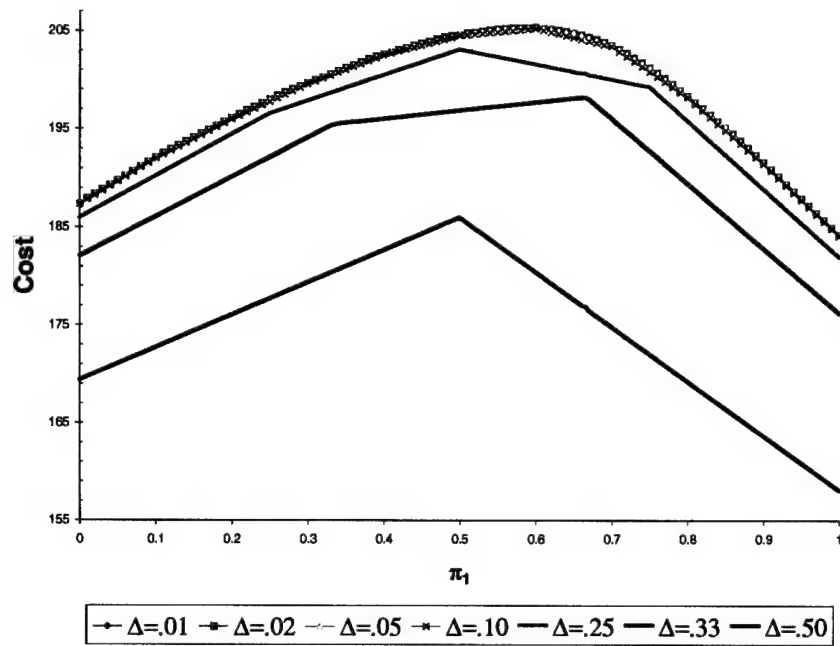
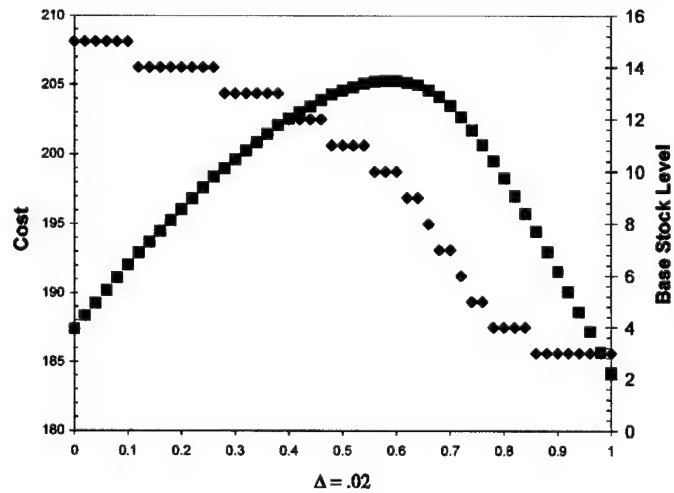
Table 4.2: Demand Distributions

Discount	=	.9
Leadtime	=	0
Critical Ratio	=	0.727
Assume observed lost sales		

Table 4.3: Experiment Parameters

4.5 Example Case

In this section we present an example of how to determine an approximation using the grid technique. We also show the subgradients that form an upper bound to the optimal cost function. There are two distributions that form the set of possible solutions. Each of the distributions were initially generated using negative binomial distributions and then truncated at a maximum demand of 18 and rescaled. Table 4.2 shows the demand parameters. The second and third columns show the mean and standard deviation before truncation, and the fourth column shows the mean of the truncated distributions.

Figure 4.3: Approximate Costs at π_1 for Different SpacingsFigure 4.4: Approximate Costs and Order Quantities at π_1

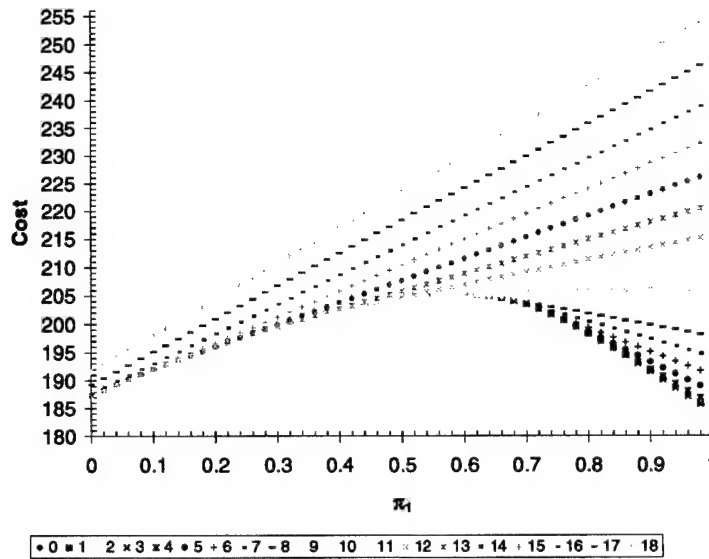


Figure 4.5: Cost-to-go function at various inventory levels

The transition matrix is

$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}.$$

The number of grid points determines the accuracy of the approximation. In Figure 4.3 we show a cost function with 7 different levels of grid spacing, $\Delta = \frac{1}{G}$. This example has two distributions, so the x-axis represents the value of π_1 . For $\Delta = .25, .33$, and $.50$, we show the linear approximation of the cost function. Notice that the cost function appears to converge very quickly once the grid size is reduced to $.10$ and smaller. This technique can serve as the actual approximation of the cost function. For example, the same case is shown in Figure 4.4 with $\Delta = .02$. The right vertical axis shows the

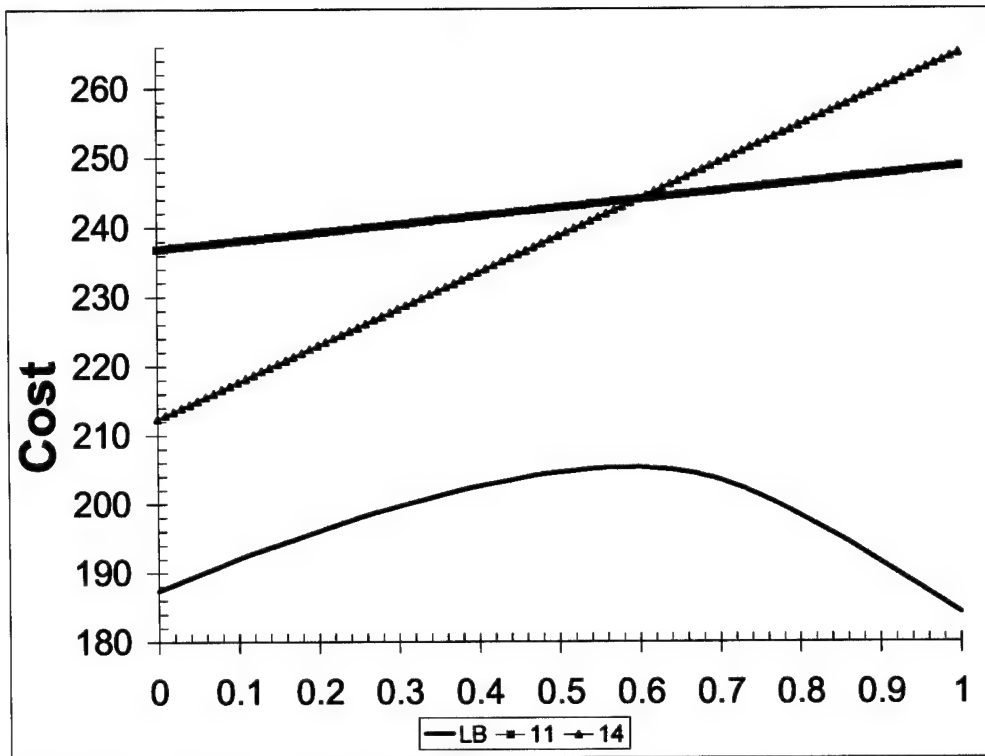


Figure 4.6: Initial Upper Bound

order-up-to quantities (ranging from 3 - 15) at each allowable value of π_1 . In Figure 4.5 we show the cost-to-go function at all possible inventory levels. A much higher penalty is incurred when the inventory level is high and there is a strong belief (π_1 close to 1) that the current demand distribution is the low demand.

In Figure 4.6 we show the optimal cost function from the grid approximation. We also show the first two subgradients that form the upper bound on this function. The first two subgradients were formed by evaluating the costs at the extreme points. The initial two subgradients were formed with static policies of 11 and 14 respectively. In Figure 4.7 we show five additional subgradients which were found by performing single

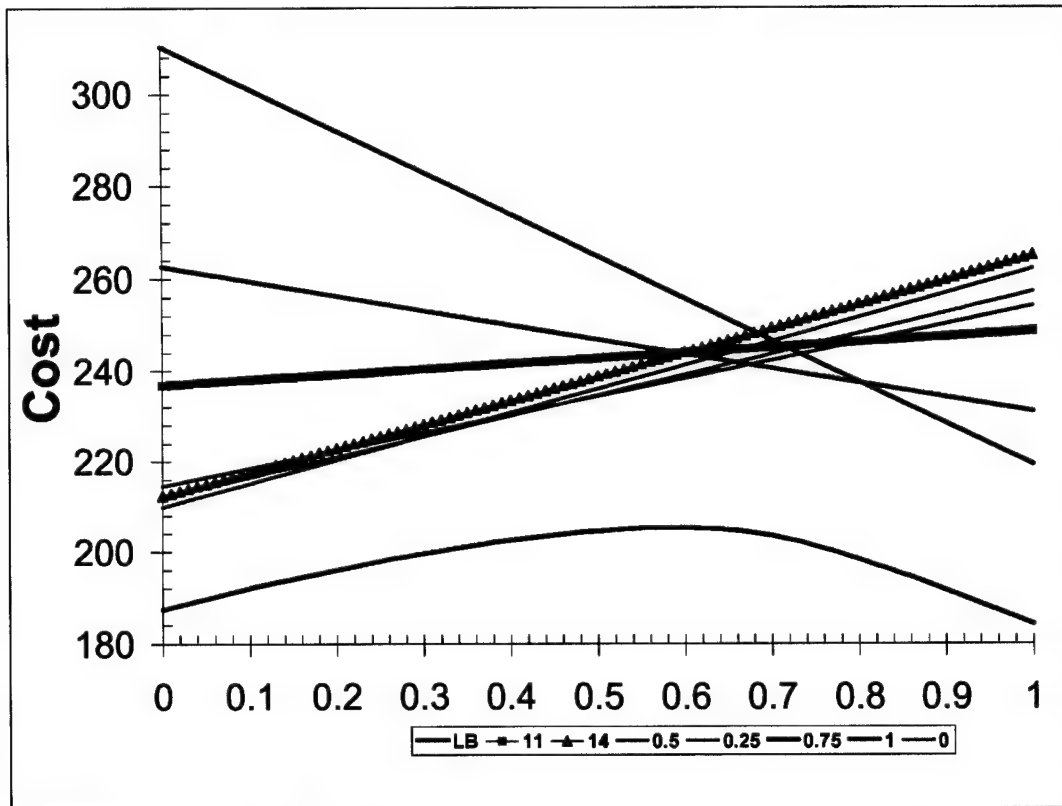


Figure 4.7: Upper Bound Improvement

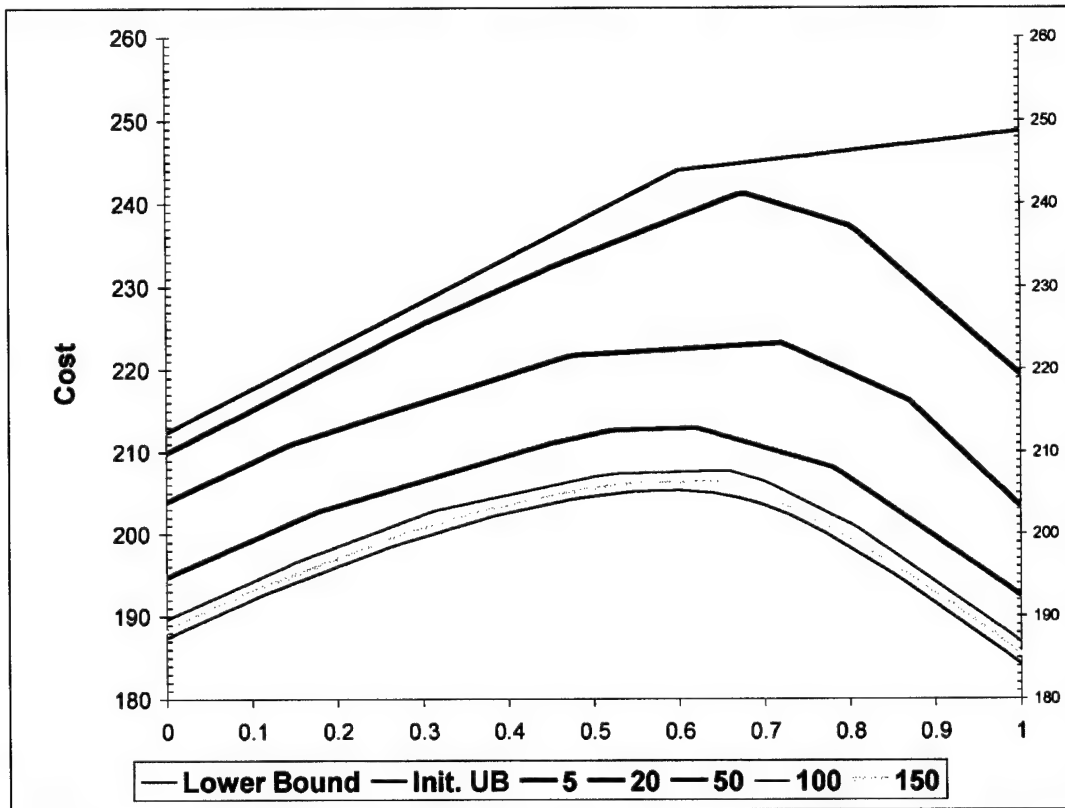


Figure 4.8: Upper Bound Convergence

Subgradient	Recursion Point (π_1)	Cost at $\pi_1 = 1$ (Low demand)	Cost at $\pi_1 = 0$ (High Demand)	Action
1	1.0	248.71	236.88	11
2	0.0	265.11	212.39	14
3	0.5	254.17	214.56	13
4	0.25	257.22	211.98	14
5	0.75	231.02	262.65	5
6	1.0	219.37	310.35	3
7	0.0	262.12	209.91	15

Table 4.4: Initial Subgradients and Actions

Lower Bound							
Grid Spacing	.50	.33	.25	.10	.05	.02	.01
Run Time	0.2	0.5	0.8	4.1	10.7	1:12	4:46
Min	9.09	2.61	0.72	0.04	0.01	0.01	0.01
Max	14.70	4.63	2.15	0.43	0.25	0.19	0.19
Avg	11.72	3.60	1.18	0.10	0.04	0.02	0.02

Table 4.5: Lower Bound Run Times and Percent Deviation from UB

dynamic recursions at $\pi_1 = .5, .25, .75, 1.0$, and 0.0 . Table 4.4 shows the values of the first seven subgradients and their respective order-up-to levels. If we perform a recursion at $\pi_1 = .5$, the resulting subgradient is based on an action of 13. We would expect it to be in the range between 11 and 14. If we are fairly confident that the demand is high ($\pi_1 = .25$), then we would expect the next subgradient to be based on a high stock level which coincides to its actual value of 14. However, if the demand is thought to be low, we order-up-to 5. If we are certain the demand is low, the initial stock level is 3. We can continue to add subgradients in this manner to improve the upper bound.

Upper Bound										
Recursions	0	5	20	50	100	150	200	250	300	500
Run Time	0.7	0.9	1.6	3.5	7.8	14.2	21.9	34.0	43.5	1:46.3
Min	13.32	12.00	8.20	3.38	1.07	0.41	0.18	0.08	0.04	.01
Max	35.01	19.73	11.59	4.46	1.53	0.79	0.41	0.29	0.23	.19
Avg	19.01	15.41	9.20	3.79	1.26	0.53	0.24	0.11	0.06	.02

Table 4.6: Upper Bound Run Times and Percent Deviation from LB

Tables 4.5 and 4.6 shows the run times on a SUN workstation to calculate a lower bound using grid approximation and an upper bound using the subgradient techniques. Table 4.5 shows the deviation from the upper bound (500 recursions) of the lower bound at various grid spacings. At a grid spacing of .10 and smaller the cost function converges very rapidly. The solution at this spacing is computed in only 4.1 seconds. Table 4.5 shows the deviation of each upper bound from the grid approximation from the best lower bound (spacing = .01). Notice that after 150 recursions that the percent deviation is less than one percent at every grid point. This is accomplished in 14.2 user seconds. This suggests that the upper bound subgradient technique can be a very valuable and efficient procedure to get a tight upper bound on the cost function.

4.6 Conclusion

In practice, demand is frequently non-stationary and also characterized by partial information. These two characteristics significantly complicate the problem. Therefore, many practitioners simplify the problem by assuming either that the demand is stationary or that the underlying demand distribution is completely known. Neither of

these assumptions is necessary. We show that the optimal cost function is pwl-cc. This characteristic allows us to use a grid approximation method to estimate the optimal cost function. This approximation may be computationally difficult to solve if the state space is large. We show two other techniques to compute both an upper and a lower bound to the optimal cost function. These bounds are not as tight as the bounds generated in a stationary problem. However, we may iteratively improve both the upper and lower bounds. These bounds can be used to produce policies that provide significant cost savings to the decision maker. Finally, we present myopic and look-ahead policies which also can be used to approximate the cost function.

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH

Our goal was to show that problems with non-stationary demand and partial information can be solved by efficient and effective procedures without making commonly used and costly assumptions. We have done so by first reviewing key literature and then presenting a basic model and reviewing optimal policies and their characteristics for both censored and uncensored stationary demand with partial information. We also discuss cases in which state space reductions make problems much easier to solve. We also note that most realistic cases will require some type of suboptimal policy. We then presented a model for the case of non-stationary demand with partial information. We model the problem as a composite-state, partially observed Markov decision process. When policies use feedback and consider all the demand uncertainty over some period of time, performance which clearly out-matches standard CEC approaches can easily be obtained. We provide limited look-ahead and open-loop feedback control models and test these policies against standard CEC approaches as well as optimal approaches. Next, we presented the same non-stationary demand problem, although with an infinite planning horizon. In this case, the optimal cost function will usually not be computable due to the intractability of the state space. Therefore, we develop a grid approximation technique that is valid for this problem. We then show how to calculate and improve both an upper and a lower bound for the optimal cost function.

Non-stationary demand is common in many production control environments. Additionally, despite a wealth of information, these same environments will often be characterized by partial information for the demand state. Managers can solve many of these problems by techniques presented in this dissertation and obtain significant monetary savings over policies commonly used today.

This research can be expanded in several important ways. One of the first areas that we would like to expand this research to is what we call constrained inventory management for a volatile demand process. This problem has non-stationary demand and partial information. However, we will also incorporate a production constraint in the system, which exists in many real life applications. We plan to provide an optimal model and a detailed discussion of its characteristics. Because optimal policies are computable for only small problems, we will also develop suboptimal policies. We will test these policies over a wide range of problem instances and compare them with an optimal policy. Finally, we will show how these procedures can provide large cost savings in a real world example.

In this dissertation, we developed models for some fairly general problems. Some of the analytical results apply to problems with various assumptions. We would like to conduct more computational testing to more thoroughly understand the nature of these problems. For example, we would like to do more computational testing with positive leadtimes and fixed setup costs. We would also like to more closely examine cases with lost sales in which the demand may not be completely observed. Lost sales problems will

lead to the concept of dual control which balances cost minimization with the value of uncensored observations. Much of this research is dependent on an externally supplied prior on the unknown parameter. We would like to conduct sensitivity analysis on the prior. This will provide important insights into the value of present information.

We would like to do more thorough testing of various transition structures. We have the capability to model increasing, decreasing, cyclical, and other demand structures. More computational testing of these situations may offer valuable insights.

Finally, a major but important extension of this research would be to extend this work to a multi-echelon problem. This extension will give important information relevant to supply chain management and control.

BIBLIOGRAPHY

- [1] Narendra Agrawal and Stephen A. Smith. Estimating Negative Binomial Demand for Retail Inventory Management with Unobservable Lost Sales. *Naval Research Logistics*, 43(6):839 – 861, September 1996.
- [2] Ravi Anupindi, Thomas E. Morton, and David Pentico. The Nonstationary Stochastic Lead-Time Inventory Problem: Near-Myopic Bounds, Heuristics, and Testing. *Management Science*, 42(1):124 – 129, 1996.
- [3] Kenneth J. Arrow, Theodore Harris, and Jacob Marshak. Optimal Inventory Theory. *Econometrica*, 19(3):250 – 272, July 1951.
- [4] Katy S. Azoury. Bayes Solution to Dynamic Inventory Models under Unknown Demand Distributions. *Management Science*, 31(9):1150 – 1160, 1985.
- [5] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 1995.
- [6] David J. Braden and Marshall Freimer. Informational Dynamics of Censored Observations. *Management Science*, 37(11):1390 – 1404, November 1991.
- [7] H. Cheng. *Algorithms for Partially Observed Markov Decision Processes*. PhD thesis, Faculty of Commerce and Business Administration, University of British Columbia, 1988.
- [8] Morris H. DeGroot. *Optimal Statistical Decisions*. McGraw-Hill, 1970.
- [9] H. Freudenthal. Simplizialzerlegungen von Beschränkter Elachheit. *Annals of Mathematics*, 43:580 – 582, 1942.
- [10] Guillermo Gallego, Jennifer K. Ryan, and David Simchi-Levi. Minimax Analysis for the Discrete Finite Horizon Inventory Model. Working paper, Columbia University, October 1996.
- [11] Srinagesh Gavirneni and Sridhar Tayur. An Efficient Procedure for Computing Optimal Order Upto Levels for Non-stationary Inventory Control. Working paper, Austin Product Center - Research, November 1998.
- [12] Stephen C. Graves. A Single-Item Inventory Model for a Non-Stationary Demand Process. Working paper, Massachusetts Institute of Technology, January 1997.

- [13] G. Hadley and T.M. Whitin. An Optimal Final Inventory Model. *Management Science*, 7:179 – 183, 1961.
- [14] W. J. Hopp, J. C. Bean, and R. L. Smith. A New Optimality Criterion for Nonhomogeneous Markov Decision Processes. *Operations Research*, 35:875 – 883, 1987.
- [15] D. L. Iglehart. The Dynamic Inventory Problem with Unknown Demand Distribution. *Management Science*, 10:429 – 440, 1964.
- [16] Alan J. Kaplan. Bayesian Approach to Inventory Control of New Parts. *IIE Transactions*, 20(2):151 – 156, June 1988.
- [17] Samuel Karlin. Dynamic Inventory Policy with Varying Stochastic Demands. *Management Science*, 6(3):231 – 258, April 1960.
- [18] Uday S. Karmarkar. A Robust Forecasting Technique for Inventory and Leadtime Management. *Journal of Operations Management*, 12:45 – 54, 1994.
- [19] Abbas A. Kurawarwala and Hirofumi Matsuo. Forecasting and Inventory Management of Short Life-Cycle Products. *Operations Research*, 44(1):131 – 150, 1996.
- [20] Martin A. Lariviere and Evan L. Porteus. Informational Dynamics and New Product Pricing. Working paper, Stanford University, March 1995.
- [21] Martin A. Lariviere and Evan L. Porteus. Stalking Information: Bayesian Inventory Management with Unobserved Lost Sales. Working paper, Fuqua School of Business, Duke University, 1997.
- [22] Hau L. Lee, V. Padmanabhan, and Seungjin Whang. Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science*, 43(4):546 – 558, 1997.
- [23] H.L. Lee and S. Nahmias. Single Product, Single-Location Models. In S.C. Graves, A.H.G. Rinnooy Kan, and P.H. Zipkin, editors, *Handbooks in Operations Research and Management Science: Logistics of Production and Inventory*, volume 4, chapter 1, pages 1 – 55. Elsevier Science, 1993.
- [24] William S. Lovejoy. Some Monotonicity Results for Partially Observed Markov Decision Processes. *Operations Research*, 35(5):736 – 743, 1987.
- [25] William S. Lovejoy. Myopic Policies for Some Inventory Models with Uncertain Demand Distributions. *Management Science*, 36(6):724 – 738, June 1990.
- [26] William S. Lovejoy. Computationally Feasible Bounds for Partially Observed Markov Decision Processes. *Operations Research*, 39:162 – 174, 1991.
- [27] William S. Lovejoy. A Survey of Algorithmic Methods for Partially Observed Markov Decision Processes. *Annals of Operations Research*, 28:47 – 65, 1991.

- [28] William S. Lovejoy. Suboptimal Policies, with Bounds, for Parameter Adaptive Decision Processes. *Operations Research*, 41(3):583 – 599, 1993.
- [29] George E. Monahan. A Survey of Partially Observable Markov Decision Processes: Theory, Models, and Algorithms. *Management Science*, 28(1):1 – 16, January 1982.
- [30] Steven Nahmias. Demand Estimation in Lost Sales Inventory Systems. *Naval Research Logistics*, 1994. To Appear.
- [31] William P. Pierskalla. An Inventory Problem with Obsolescence. *Naval Research Logistics Quarterly*, 16:217 – 228, 1969.
- [32] D. Rhenius. Incomplete Information in Markovian Decision Models. *Annals of Statistics*, 2:1327 – 1334, 1974.
- [33] Herbert E. Scarf. Bayes Solution of the Statistical Inventory Problem. *Annals of Mathematical Statistics*, 30:490 – 508, 1959.
- [34] Herbert E. Scarf. Some Remarks on Bayes Solution to the Inventory Problem. *Naval Research Logistics Quarterly*, 7:591 – 596, 1960.
- [35] Richard D. Smallwood and Edward J. Sondik. The Optimal Control of Partially Observable Markov Processes over a Finite Horizon. *Operations Research*, 21:1071 – 1088, 1973.
- [36] Edward J. Sondik. The Optimal Control of Partially Observable Markov Processes over the Infinite Horizon: Discounted Costs. *Operations Research*, 26(2):282 – 304, 1978.
- [37] Jing-Sheng Song and Paul Zipkin. Inventory Control in a Fluctuating Demand Environment. *Operations Research*, 41(2):351 – 370, 1993.
- [38] Jing-Sheng Song and Paul H. Zipkin. Managing Inventory with the Prospect of Obsolescence. *Operations Research*, 44(1):215 – 222, 1996.
- [39] James T. Treharne and Charles R. Sox. Adaptive Inventory Control for Stationary Demand and Partial Information. *Working Paper*, Auburn University, September 1998.
- [40] James T. Treharne and Charles R. Sox. Adaptive Inventory Control for Non-Stationary Demand and Partial information. *Working Paper*, Auburn University, January 1999.
- [41] K.M. Van Hee. Bayesian Control of Markov Chains. In *Mathematical Centre Tract 95*, Amsterdam, 1978.

- [42] Arthur F. Veinott. Optimal Stockage Policies with Non-Stationary Stochastic Demands. In Herbert E. Scarf, Dorothy M. Gilford, and Maynard W. Shelly, editors, *Multistage Inventory Models and Techniques*, Office of Naval Research Monograph Series on Mathematical Methods in Logistics, chapter 4, pages 85 – 115. Stanford University Press, 1963.
- [43] Chelsea C. White and William T. Scherer. Solution Procedures and Partially Observed Markov Decision Processes. *Operations Research*, 37(5):791 – 797, 1989.
- [44] Chelsea C. White and William T. Scherer. Finite-Memory Suboptimal Design for Partially Observed Markov Decision Processes. *Operations Research*, 42(3):439 – 455, 1994.
- [45] Chelsea C. White III and Douglass J. White. Markov Decision Processes. *European Journal of Operational Research*, 39:1– 16, 1989.